

THE ERROR TERM OF THE PRIME ORBIT THEOREM FOR EXPANDING SEMIFLOWS

MASATO TSUJII

ABSTRACT. We consider suspension semiflows of angle multiplying maps on the circle and study the distributions of periods of their periodic orbits. Under generic conditions on the roof function, we give an asymptotic formula on the number $\pi(T)$ of prime periodic orbits with period $\leq T$. The error term is bounded, at least, by

$$\exp \left(\left(1 - \frac{1}{4\lceil \chi_{\max}/h_{\text{top}} \rceil} + \varepsilon \right) h_{\text{top}} \cdot T \right) \quad \text{in the limit } T \rightarrow \infty$$

for arbitrarily small $\varepsilon > 0$, where h_{top} and χ_{\max} are respectively the topological entropy and the maximal Lyapunov exponent of the semiflow.

1. INTRODUCTION

For a flow $f^t : M \rightarrow M$ on a closed manifold M with some hyperbolicity, it is well known that the number $\pi(T)$ of periodic orbits with period $\leq T$ grows exponentially as $T \rightarrow \infty$ and the exponential rate coincides with the topological entropy h_{top} of the flow. The prime orbit theorem, due to Parry and Pollicott [10, Theorem 9.3], gives a more precise estimate in the case of topologically weakly mixing hyperbolic flows:

$$(1) \quad \pi(T) = (1 + o(1)) \int_1^T \frac{e^{h_{\text{top}} t}}{t} dt \quad \text{as } T \rightarrow \infty.$$

This paper addresses estimates of the error term in this asymptotic formula.

For geodesic flows on surfaces with negative (variable) curvature, Pollicott and Sharp proved in [12] that the relative error term, denoted by $o(1)$ in the formula (1) above, is actually exponentially small, that is, bounded by $Ce^{-\varepsilon T}$ with some $C > 0$ and $\varepsilon > 0$. More recently, this result is extended to higher dimensional cases by Giulietti, Liverani and Pollicott[3] and Stoyanov[14]. But not much is known about the exponential rate at which the relative error term decreases.

For the geodesic flows on surfaces with negative *constant* curvature, we have a much more precise asymptotic formula due to Huber, which reads

$$(2) \quad \pi(T) = \int_1^T \frac{e^{h_{\text{top}} t}}{t} dt + \sum_{i=1}^k \int_1^T \frac{e^{\mu_i t}}{t} dt + \mathcal{O}(e^{\rho t})$$

where $\rho = (3/4)h_{\text{top}}$ and μ_i , $1 \leq i \leq k$, are real numbers satisfying $\rho < \mu_i < h_{\text{top}}$. (The exponents μ_i are related to small eigenvalues of the Laplacian on the surface. See [2].) But this result is known only for the case of constant curvature because the proof is based on the fact that the geodesic flow in such case is identified with

the left action of a hyperbolic one-parameter subgroup of $SL(2, \mathbb{R})$ on its quotient space by the right action of a discrete subgroup.

Comparing these results, we are tempted to pose a question whether such a precise asymptotic formula as (2) is available for more general type of hyperbolic flows and by a more flexible method. In this paper, we pursue this question in the case of suspension semiflows of angle multiplying maps on the circle and provide a positive answer under generic conditions on the roof function.

2. THE MAIN RESULTS

2.1. Definitions. We consider a class of (simplest possible) expanding semiflows. This kind of semiflows have been studied in [13, 11, 17] as a simplified model of Anosov flows. First we fix a positive integer $\ell \geq 2$ and consider the angle-multiplying map

$$\tau : S^1 \rightarrow S^1, \quad \tau(x) = \ell x \mod \mathbb{Z}.$$

Let $C_+^\infty(S^1)$ be the space of positive-valued C^∞ functions on S^1 . Then we consider the suspension semiflow of τ with roof function $f \in C_+^\infty(S^1)$:

$$\mathbf{T}_f = \{T_f^t : X_f \rightarrow X_f \mid t \geq 0\}.$$

(See Figure 1.) This is a semiflow on the set

$$X_f := \{(x, y) \in S^1 \times \mathbb{R} \mid 0 \leq y < f(x)\} \subset S^1 \times \mathbb{R}$$

and defined precisely by the expression

$$T_f^t(x, y) = (\tau^{n(x, y+t; f)}(x), y + t - f^{(n(x, y+t; f))}(x))$$

where

$$(3) \quad f^{(n)}(x) = \sum_{i=0}^{n-1} f(\tau^i(x))$$

and

$$(4) \quad n(x, t; f) = \max\{n \geq 0 \mid f^{(n)}(x) \leq t\}.$$

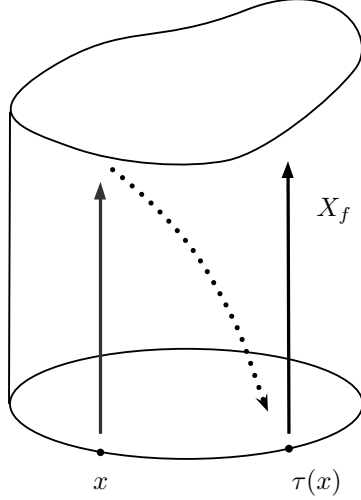
2.2. Spectral properties of transfer operators. By a heuristic argument, the distribution of periods of periodic orbits of \mathbf{T}_f is related to the spectra of the transfer operators

$$\mathcal{L}^t \varphi(z) = \sum_{w: T_f^t(w)=z} \varphi(w).$$

Indeed, computing the flat trace of \mathcal{L}^t , defined as the integral of the Schwartz kernel $K^t(z, w)$ of \mathcal{L}^t along the diagonal $z = w$, we find

$$(5) \quad \text{Tr}^b \mathcal{L}^t = \sum_{\gamma \in \Gamma} \sum_{n=1}^{\infty} \frac{|\gamma|}{1 - E_\gamma^{-n}} \cdot \delta(t - n|\gamma|)$$

where Γ is the set of prime periodic orbits and $|\gamma|$ and E_γ denote respectively the prime period and the (coefficient of) linearized Poincaré map. If we ignore the sum

FIGURE 1. Expanding semiflow \mathbf{T}_f

over $n \geq 2$ and also the term E_γ^{-n} in the denominator of the summands (which are in fact relatively small), we would have

$$(6) \quad \frac{1}{t} \cdot \text{Tr}^b \mathcal{L}^t \sim \sum_{\gamma \in \Gamma} \delta(t - |\gamma|), \quad \text{and so} \quad \int_1^T \frac{1}{t} \cdot \text{Tr}^b \mathcal{L}^t dt \sim \pi(T).$$

Therefore, if the flat trace $\text{Tr}^b \mathcal{L}^t$ were related to the spectrum of \mathcal{L}^t as in the case of the usual trace, the asymptotics of $\pi(T)$ would be expressed in terms of the spectrum of \mathcal{L}^t . For this reason, we are going to study the spectral properties of the transfer operators \mathcal{L}^t .

Let us say that a function $\varphi : X_f \rightarrow \mathbb{C}$ is of class C^∞ if $\mathcal{L}^t \varphi$ for $t \geq 0$ are C^∞ functions on the interior X_f° of X_f (as a subset of $S^1 \times \mathbb{R}$) and each of their partial derivatives are bounded. This implies that φ is not only C^∞ on X_f° but also satisfies some conditions on its values and differentials on the boundaries. Let $C^\infty(X_f)$ be the space of C^∞ functions on X_f (defined as above), which, equipped with the uniform C^r norms $\|\varphi|_{X_f^\circ}\|_{C^r}$ for $r \geq 0$, is a Fréchet space. With this definition of $C^\infty(X_f)$, we may regard \mathcal{L}^t for $t \geq 0$ as a continuous semigroup of operators

$$\mathcal{L}^t : C^\infty(X_f) \rightarrow C^\infty(X_f).$$

To study spectral properties of \mathcal{L}^t , we will define Banach spaces

$$C^\infty(X_f) \subset \mathcal{B}^{r,p}(X_f) \subset L^2(X_f)$$

for real numbers $r > 0$ and integers $p \geq 1$ and consider the natural extensions of \mathcal{L}^t to them. The next theorem gives a spectral property of \mathcal{L}^t on $\mathcal{B}^{r,p}(X_f)$ under some generic conditions on the roof function f . We write $h(f)$, $\chi_{\max}(f)$ and $\chi_{\min}(f)$ respectively for the topological entropy, the maximum Lyapunov exponent and the minimum Lyapunov exponent:

$$\chi_{\max}(f) := \lim_{t \rightarrow \infty} \frac{1}{t} \max_{z \in X_f} \log \|DT_f^t(z)\|, \quad \chi_{\min}(f) := \lim_{t \rightarrow \infty} \frac{1}{t} \min_{z \in X_f} \log \|DT_f^t(z)\|.$$

We put

$$\alpha(f) := \frac{\chi_{\max}(f)}{h(f)}.$$

We always have $\alpha(f) \geq 1$ from Ruelle inequality[8] and may regard $\alpha(f)$ as a measurement of spacial non-uniformity of expansion by the semiflow \mathbf{T}_f .

Theorem 2.1. *For any $f \in C_+^\infty(S^1)$, any $r > 0$ and any integer $p \geq 1$, the transfer operators \mathcal{L}^t for sufficiently large $t > 0$ extend to bounded operators*

$$(7) \quad \mathcal{L}^t : \mathcal{B}^{r,p}(X_f) \rightarrow \mathcal{B}^{r,p}(X_f).$$

For each integer $p \geq 1$ and for each $\varepsilon > 0$, there is an open dense subset $\mathcal{U}_p(\varepsilon) \subset C_+^\infty(S^1)$ such that, if $f \in \mathcal{U}_p(\varepsilon)$ and if $r > 0$ is sufficiently large (see (16)), the essential spectral radius of the operator (7) for sufficiently large $t > 0$ is smaller than $\exp((\rho_p(f) + \varepsilon)t)$ where

$$(8) \quad \rho_p(f) := \frac{1}{2} \left(1 + \frac{\max\{p, \alpha(f)\} - 1}{p} \right) \cdot h(f).$$

Remark 2.2. The conclusion of the theorem above implies that the spectral set of (7) on the region $|z| \geq \exp((\rho_p(f) + \varepsilon)t)$ consists of finitely many eigenvalues with finite multiplicities. Such eigenvalues (counted with multiplicity) are written in the form $\exp(\mu_i t)$, $i = 1, 2, \dots, k$, with complex numbers μ_i that do not depend on t . (See [17, pp295].)

The case $p = 1$ in the theorem above corresponds to the result in our previous paper [17], where the bound is

$$\rho_1(f) = \exp(\chi_{\max}(f) \cdot t/2)$$

as $\alpha(f) \geq 1$. (See also [18, 19] for the corresponding results for contact Anosov flows.) This bound is preferable when $\alpha(f)$ is close to 1, but the claim becomes vacuous when $\alpha(f) \geq 2$ for $\rho_1(f)$ exceeds the topological entropy $h_{\text{top}}(f)$. The improvement achieved in Theorem 2.1 is that we get better bounds by choosing different integers $p \geq 1$ depending on $\alpha(f) \geq 1$. For simplicity's sake, suppose that f belongs to the residual subset $\mathcal{U} := \cap_{p \in \mathbb{N}} \cap_{m=1}^\infty \mathcal{U}_p(1/m) \subset C_+^\infty(S^1)$ and put

$$\rho(f) := \min_{p \geq 1} \rho_p(f).$$

Letting ¹ $p(f) = \lceil \alpha(f) \rceil \geq 1$, we have

$$(9) \quad \rho(f) \leq \rho_{p(f)}(f) \leq \left(1 - \frac{1}{2p(f)} \right) h(f) < h(f).$$

Therefore, by choosing suitable $p \geq 1$, we always get a bound for the essential spectral radius of \mathcal{L}^t that is strictly smaller than the spectral radius $\exp(h(f)t)$.

2.3. Asymptotics of the number of periodic orbits. We next give a consequence of Theorem 2.1 on the remainder term of the prime orbit theorem. Let $\Gamma = \Gamma(f)$ be the set of prime periodic orbits for the semiflow \mathbf{T}_f . For a prime periodic orbit $\gamma \in \Gamma$, we denote its period by $|\gamma|$. Let $\pi(T) = \#\{\gamma \in \Gamma \mid |\gamma| \leq T\}$.

¹This choice of p is not always optimal.

Theorem 2.3. *Let $\varepsilon > 0$ and suppose that the roof function $f \in C_+^\infty(S^1)$ belongs to the open dense subset $\mathcal{U}_p(\varepsilon) \subset C_+^\infty(X_f)$ given in Theorem 2.1 for $p \geq 1$. Then, with setting*

$$(10) \quad \bar{\rho} = \bar{\rho}_p(f) := \frac{\rho_p(f) + h(f)}{2},$$

we have an asymptotic formula

$$\pi(T) = \int_1^T \frac{e^{h_{\text{top}} t}}{t} dt + \sum_{i=1}^{k'} \int_1^T \frac{e^{\mu_i t}}{t} dt + \mathcal{O}\left(e^{(\bar{\rho} + \varepsilon)t}\right),$$

where μ_i , $1 \leq i \leq k'$, are those complex numbers in Remark 2.2 satisfying the condition $\bar{\rho} + \varepsilon < \Re(\mu_i) < h(f)$.

Note that, if we let $p = p(f) = \lceil \alpha(f) \rceil \geq 1$, we have, from (9), that

$$\bar{\rho}_p(f) \leq \left(1 - \frac{1}{4 \lceil \chi_{\max}(f)/h(f) \rceil}\right) h(f) < h(f).$$

Remark 2.4. The reason that we have the average $\bar{\rho}_p(f)$ instead of $\rho_p(f)$ in the statement of Theorem 2.3 above will appear in its proof given in Section 6. (See Remark 6.9.)

In the following sections, we proceed as follows. In Section 3, we formulate a transversality condition on the roof function f and decompose Theorem 2.1 into two theorems: Theorem 3.3 that proves prevalence of the transversality condition and Theorem 3.5 that proves the conclusion of Theorem 2.1 from the transversality condition. We prove Theorem 3.5 in Section 5 after preparation in Section 4. We insert the proof of Theorem 2.3 in Section 6 before we prove Theorem 3.3 in Section 7, as it uses the argument in the proof of Theorem 3.5.

3. THE GENERIC CONDITION

We set up notation on the dynamics of the semiflow \mathbf{T}_f and formulate the transversality condition that defines the open dense subset $\mathcal{U}_p(\varepsilon)$ in Theorem 2.1.

3.1. Differential of the semiflow \mathbf{T}_f . The differential $DT_f^t(z) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ at $z \in X_f$ is well-defined if z and $T_f^t(z)$ are not on the (lower) boundary of X_f . In general, we define

$$DT_f^t(z) = \lim_{\varepsilon \rightarrow +0} DT_f^t(x, y + \varepsilon) : \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

where $DT_f^t(x, y + \varepsilon)$ for sufficiently small $\varepsilon > 0$ is constant and hence the limit on the right hand side is well-defined. For $t \geq 0$, we set

$$(11) \quad E(z, t; f) = \ell^{n(x, y+t; f)} \quad \text{and} \quad F(z, t; f) = \frac{\partial}{\partial x} f^{(n(x, y+t; f))}(x)$$

where $n(x, t; f)$ and $f^{(n)}(x)$ are those defined in (3) and (4). Then we have

$$(12) \quad DT_f^t(z) = \begin{pmatrix} E(z, t; f) & 0 \\ F(z, t; f) & 1 \end{pmatrix}$$

We write $D^\dagger T_f^t(z)$ for the transpose of the inverse of $Df^t(z)$, that is,

$$(13) \quad D^\dagger T_f^t(z) := {}^T(Df^t(z))^{-1} = \begin{pmatrix} E(z, t; f)^{-1} & S(z, t; f) \\ 0 & 1 \end{pmatrix}$$

where

$$(14) \quad S(z, t; f) = -E(z, t; f)^{-1}F(z, t; f).$$

Then the minimum and maximum Lyapunov exponents of \mathbf{T}_f are written

$$\chi_{\min}(f) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \left(\min_{z \in X_f} E(z, t; f) \right)$$

and

$$\chi_{\max}(f) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \left(\max_{z \in X_f} E(z, t; f) \right).$$

For the topological entropy $h(f)$, we have

$$\frac{1}{t} \log \left(\min_{z \in X_f} E(z, t; f) \right) \leq h(f) \leq \frac{1}{t} \log \left(\max_{z \in X_f} E(z, t; f) \right)$$

for any $t > 0$ and hence

$$\chi_{\min}(f) \leq h(f) \leq \chi_{\max}(f).$$

For $0 < y_{\min} < y_{\max}$ and $\kappa_0 > 0$, let $\mathfrak{F}(y_{\min}, y_{\max}, \kappa_0) \subset C_+^\infty(S^1)$ be the open subset that consists of $f \in C_+^\infty(S^1)$ satisfying

$$y_{\min} < f(x) < y_{\max}, \quad |f'(x)| < \kappa_0, \quad |f''(x)| < \kappa_0 \quad \text{for all } x \in S^1.$$

If $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$, we have

$$(15) \quad \bar{\chi}_{\min} := \frac{\log \ell}{y_{\max}} \leq \chi_{\min}(f) \leq h(f) \leq \chi_{\max}(f) \leq \bar{\chi}_{\max} := \frac{\log \ell}{y_{\min}}.$$

In what follows, we fix $0 < y_{\min} < y_{\max}$ and $\kappa_0 > 0$ and confine our attention to the semiflows \mathbf{T}_f with $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$. Since the subset $\mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$ exhausts $C_+^\infty(S^1)$ in the limit $y_{\min} \rightarrow +0$, $y_{\max} \rightarrow +\infty$ and $\kappa_0 \rightarrow +\infty$, this causes no loss of generality. We henceforth fix $r > 0$ such that

$$(16) \quad r > \bar{\chi}_{\max}/\bar{\chi}_{\min} \geq 1.$$

3.2. Cones in the flow direction. Since the time- t -map T_f^t is partially hyperbolic, its (push-forward) action on the cotangent bundle

$$D^\dagger T_f^t : X_f \times \mathbb{R}^2 \rightarrow X_f \times \mathbb{R}^2, \quad D^\dagger T_f^t(z, \xi) = (T_f^t(z), D^\dagger T_f^t(z)\xi)$$

admits a forward invariant cone field. We can set up such a cone field concretely as follows. For real numbers s and $\theta > 0$, we define

$$\mathbf{C}(s, \theta) := \{(\xi, \eta) \in \mathbb{R}^2 \mid |\xi - s\eta| \leq \theta|\eta|\} \subset \mathbb{R}^2.$$

We fix a real number γ_0 satisfying $1/\ell < \gamma_0 < 1$ and set

$$\mathbf{C}_0 := \mathbf{C}(0, \theta_0) := \{(\xi, \eta) \in \mathbb{R}^2 \mid |\xi| \leq \theta_0|\eta|\} \subset \mathbb{R}^2$$

where

$$\theta_0 := \frac{\kappa_0}{\gamma_0 \ell - 1}.$$

Then we have that

$$(17) \quad (DT_f^t)_z^\dagger(\mathbf{C}_0) = \mathbf{C}(S(z, t; f), E(z, t; f)^{-1}\theta_0) \subset \mathbf{C}(0, \gamma_0\theta_0) \subset \mathbf{C}_0$$

for all $z = (x, y) \in X_f$ and $t \geq f(x) - y$.

3.3. Backward orbits. For each $z \in X_f$, the number of points in its backward orbit

$$(T_f^t)^{-1}(z) = \{w \in X_f \mid T_f^t(w) = z\}$$

for time $t > 0$ grows exponentially as $t \rightarrow 0$. Indeed, for any $\varepsilon > 0$, there exists $C_\varepsilon > 1$ such that

$$(18) \quad C_\varepsilon^{-1} e^{(h(f)-\varepsilon)t} < \#(T_f^t)^{-1}(z) < C_\varepsilon e^{(h(f)+\varepsilon)t} \quad \forall z \in X_f, \quad \forall t \geq 0.$$

For $z = (x, y) \in X_f$, $t \geq 0$ and $w \in (T_f^t)^{-1}(z)$, let

$$(19) \quad 0 < s_{k(z,w;t)}(z, w; t) < \cdots < s_2(z, w; t) < s_1(z, w; t) \leq t$$

be the sequence of time t at which the orbit $T_f^s(w)$, $0 < s \leq t$, crosses the lower boundary $S^1 \times \{0\}$ of X_f . By definition, we have

$$T_f^{s_k(z,w;t)}(w) \in \tau^{-k}(x) \times \{0\} \quad \text{for } 1 \leq k \leq k(z, w; t).$$

Since we are assuming that $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$, we have

$$\lfloor t/y_{\max} \rfloor \leq n(z, w; t) \leq \lceil t/y_{\min} \rceil.$$

Below we investigate transversality between the cones

$$(20) \quad (D^\dagger T_f^t)_w(\mathbf{C}_0) = \mathbf{C}(S(w, t; f), E(w, t; f)^{-1}\theta_0) \quad \text{for } w \in (T_f^t)^{-1}(z)$$

in some generalized sense. Since much variety of angles of the cones $(D^\dagger T_f^t)_w(\mathbf{C}_0)$ for $w \in (T_f^t)^{-1}(z)$ causes technical difficulties in the following argument, we classify the points $w \in (T_f^t)^{-1}(z)$ with respect to the value of $E(w, t; f)$ (whose reciprocal is proportional to the angle of $(D^\dagger T_f^t)_w(\mathbf{C}_0)$). For an interval $J = [a, b]$ with $0 < a < b$, we set

$$B(z, t; J; f) = \{w \in (T_f^t)^{-1}(z) \mid e^{at} \leq E(w, t; f) \leq e^{bt}\}.$$

We fix a C^∞ function $\chi : \mathbb{R} \rightarrow [0, 1]$ such that

$$(21) \quad \chi(t) = \begin{cases} 0, & \text{if } t \geq 2; \\ 1, & \text{if } t \leq 1. \end{cases}$$

For $s \in \mathbb{R}$, let $\langle s \rangle = \chi(s) + (1 - \chi(s))|s|$, so that $\langle s \rangle \in [1, \max\{1, |s|\}]$ and that

$$\langle s \rangle = \begin{cases} 1, & \text{if } |s| \leq 1; \\ |s|, & \text{if } |s| \geq 2. \end{cases}$$

Definition 3.1. For $z \in X_f$, $t > 0$ and a p -tuple $\mathbf{w} = (\mathbf{w}(1), \dots, \mathbf{w}(p))$ of points in $(T_f^t)^{-1}(z)$, we set

$$S(\mathbf{w}, t; f) = \sum_{i=1}^p S(\mathbf{w}(i), t; f)$$

and define $E(\mathbf{w}, t; f)$ by the relation

$$\frac{1}{E(\mathbf{w}, t; f)} = \sum_{i=1}^p \frac{1}{E(\mathbf{w}(i), t; f)}.$$

We define the function $W^r(\mathbf{w}, t; f) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by

$$W^r(\mathbf{w}, t; f)(\xi, \eta) = \left\langle \frac{E(\mathbf{w}, t; f) \cdot |\xi - S(\mathbf{w}, t; f)\eta|}{\theta_0 \cdot \langle \eta \rangle} \right\rangle^r.$$

This function takes constant value 1 on the cone

$$(22) \quad \mathbf{C}(S(\mathbf{w}, t; f), E(\mathbf{w}, t; f)^{-1}\theta_0)$$

and grows rapidly as the point gets far from it.

As a quantification of transversality between p -tuples of cones in (20), we consider the quantity

$$(23) \quad \sup_{\xi \in \mathbb{R}} \left(\sum_{\mathbf{w}=(w(1), \dots, w(p)) \in B(z, t; J; f)^p} \frac{1}{W^r(\mathbf{w}, t; f)(\xi, 2)} \right).$$

In the case $p = 1$, boundedness of this quantity by some relatively small constant implies that most of the cones in (20) are transversal to each other. In the case $p > 1$, it does not have such geometric meaning, but still useful in the argument below.

Remark 3.2. By definition, $W^r(\mathbf{w}, t; f)(\xi, \eta)$ is constant on the intersection of a straight line through the origin with the region $|\eta| \geq 2$. Hence the constant 2 in $W^r(\mathbf{w}, t; f)(\xi, 2)$ above could be any constant η_0 such that $|\eta_0| \geq 2$.

The next theorem gives a bound on (a slight modification of) the quantity (23) under generic conditions on the roof function f . Before stating the theorem, let us make a guess on the bound. Note that each function $\xi \mapsto W^r(\mathbf{w}, t; f)(\xi, 1)^{-1}$ decays rapidly on the outside of a neighborhood of $\xi = S(\mathbf{w}, t; f)$ with width proportional to $E(\mathbf{w}, t; f)^{-1} \leq e^{-at}$. Hence, if the values of $S(\mathbf{w}, t; f)$ for $\mathbf{w} \in B(z, t; J; f)^p$ were distributed randomly and independently on the interval $[-p\theta_0, p\theta_0]$ (as random variables on the space of roof functions f), the large deviation argument would tell that, for almost all roof functions f , the quantity (23) should be bounded by

$$e^{\varepsilon t} \max\{1, \exp(-at) \cdot (\sharp(T_f^t)^{-1}(z))^p\} \leq \exp((\max\{p \cdot h(f) - a, 0\} + \varepsilon)t)$$

in the limit $t \rightarrow \infty$, for arbitrarily small $\varepsilon > 0$. The next theorem tells that this guess is basically true, but with slight modifications.

For an integer $n \geq 1$, let $\text{Per}(\tau, n)$ be the set of periodic points of τ with period not greater than n and, for $\delta > 0$, let $\text{Per}_\delta(\tau, n)$ be the open δ -neighborhood of $\text{Per}(\tau, n)$.

Theorem 3.3. *Let $p \geq 1$ be an integer. For an interval $J = [a, b]$ with $0 < a < b$ and real numbers $\varepsilon, \delta > 0$, there exists $n_0 = n_0(\varepsilon)$ and a prevalent² subset*

$$\mathcal{G}(J, n, \varepsilon, \delta; p) \subset \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$$

for $n \geq n_0$, such that the following claim holds for $f \in \mathcal{G}(J, n, \varepsilon, \delta; p)$:

For sufficiently large $t > 0$ and for any $z = (x, y) \in X_f$ with $x \notin \text{Per}_\delta(n, \tau)$, there exists a subset $\mathcal{E} = \mathcal{E}(z, t; f) \subset \tau^{-n}(x)$ with $\#\mathcal{E} \leq p \lceil 10a/\varepsilon \rceil$ such that

$$(24) \quad \sum^* \frac{1}{W^r(\mathbf{w}, t; f)(\xi, 2)} < \exp((\max\{p \cdot h(f) - a, 0\} + p(b - a) + \varepsilon)t)$$

for any $\xi \in [-\theta_0, \theta_0]$, where the sum \sum^* is taken over p -tuples $\mathbf{w} = (\mathbf{w}(1), \dots, \mathbf{w}(p))$ in $B(z, t; J; f)^p$ with

$$T_f^{s_n(z, \mathbf{w}(i); t)}(\mathbf{w}(i)) \notin \mathcal{E} \times \{0\} \quad \text{for } i = 1, 2, \dots, p.$$

²See remark below for the definition of this term “prevalent”.

Note that the point $T_f^{s_n(z, \mathbf{w}(i); t)}(\mathbf{w}(i))$ belongs to $\tau^{-n}(x) \times \{0\}$ from the definition of $s_n(z, w; t)$ in (19).

Remark 3.4. In the statement above, we used the notion of “prevalence” that is introduced in [5, 6, 9]. A measurable subset S in a linear topological space X is said to be *shy* if there exists a Borel measure μ such that $0 < \mu(U) < \infty$ for some compact subset $U \subset X$ and $\mu(S + x) = 0$ for any $x \in X$. (μ is called a transverse measure for S .) A shy subset has empty interior. Any countable union of shy subsets is again shy. (This is far from trivial.) A measurable subset P is said to be *prevalent* in $Q \subset X$ if $Q \setminus P$ is shy. (See [9] for the detail. See also [16] for a similar but different notion which could be used alternatively.)

The next theorem states that the transversality condition in the theorem above yields an estimate on the essential spectral radius of the transfer operator \mathcal{L}^t .

Theorem 3.5. *Let $p \geq 1$ be an integer and let $J_\nu = [a_\nu, b_\nu]$, $1 \leq \nu \leq \nu_0$, be intervals such that the union of their interiors contains the interval $[\bar{\chi}_{\min}, \bar{\chi}_{\max}]$. For $1 \leq \nu \leq \nu_0$, put*

$$(25) \quad \mu_\nu = \frac{(p-1)h(f) + \max\{p \cdot h(f) - a_\nu, 0\} + p(b_\nu - a_\nu) + b_\nu}{2p}.$$

Suppose that f_0 belongs to the prevalent subset

$$\mathcal{G} = \bigcap_{\nu=1}^{\nu_0} \bigcap_{m=1}^{\infty} \bigcap_{m'=1}^{\infty} \bigcap_{n \geq n_0(1/m)} \mathcal{G}(J_\nu, n, 1/m, 1/m'; p) \subset \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$$

where $\mathcal{G}(J, n, \varepsilon, \delta; p)$ is that in Theorem 3.3. Then, for any $\eta > 0$, there exists a neighborhood \mathcal{V} of f_0 in $C_+^\infty(S^1)$ such that, if $f \in \mathcal{V}$, the essential spectral radius of the transfer operator (7) for sufficiently large t is bounded by $e^{(\mu(f) + \eta)t}$ where

$$(26) \quad \mu(f) = \max\{\mu_\nu \mid \text{int } J_\nu \cap [\chi_{\min}(f), \chi_{\max}(f)] \neq \emptyset\}.$$

For given $\eta > 0$, we can take the intervals $J_\nu = [a_\nu, b_\nu]$, $1 \leq \nu \leq \nu_0$, narrow enough so that the quantity $\mu(f)$ is bounded by

$$\frac{(p-1 + \max\{p, \alpha(f)\})h(f) + \eta}{2p} = \rho_p(f) + \frac{\eta}{2p}.$$

Therefore Theorem 2.1 follows from Theorem 3.3 and Theorem 3.5. (The first claim on boundedness of \mathcal{L}^t is proved in Subsection 5.1.)

4. THE BANACH SPACE $\mathcal{B}^{r,p}(\mathbb{R}^2)$

In this section, we define the Banach space $\mathcal{B}^{r,p}(\mathbb{R}^2)$ and prove some related lemmas. (The definition resembles that of Besov spaces in [15, Section 2.3.1].) We will construct the Banach space $\mathcal{B}^{r,p}(X_f)$ in (7) using this Banach space as its local model. (But, since X_f is not a manifold, the construction is a little different from the usual one.)

4.1. Definitions. We introduce two partitions of unity on \mathbb{R} :

$$\{\chi_n : \mathbb{R} \rightarrow [0, 1]\}_{n \in \mathbb{Z}_+} \quad \text{and} \quad \{\rho_n : \mathbb{R} \rightarrow [0, 1]\}_{n \in \mathbb{Z}}.$$

The former is the Littlewood-Paley partition of unity, defined by

$$\chi_m : \mathbb{R} \rightarrow [0, 1], \quad \chi_m(t) = \begin{cases} \chi(|t|), & \text{if } m = 0; \\ \chi(2^{-m}|t|) - \chi(2^{-m+1}|t|), & \text{if } m \geq 1 \end{cases}$$

where χ is that taken in Subsection 3.3. The latter is defined by

$$\rho_n = \begin{cases} \chi(\operatorname{sgn}(x)\sqrt{|x|} - n + 1) - \chi(\operatorname{sgn}(x)\sqrt{|x|} - n + 2), & \text{if } n \geq 1; \\ \chi(\sqrt{|x|} + 1), & \text{if } n = 0; \\ \chi(\operatorname{sgn}(x)\sqrt{|x|} + n + 1) - \chi(\operatorname{sgn}(x)\sqrt{|x|} + n + 2), & \text{if } n \leq -1. \end{cases}$$

Note that the support of the function ρ_n is contained in the interval

$$I_n = \begin{cases} [(n-1)^2, (n+1)^2], & \text{if } n \geq 1; \\ [-1, 1], & \text{if } n = 0; \\ [-(|n|+1)^2, -(|n|-1)^2], & \text{if } n \leq -1 \end{cases}$$

which contains $\operatorname{sgn}(n) \cdot n^2$ and whose length is comparable with $|n|$.

Next we define the partition of unity

$$\{\chi_{n,m} : \mathbb{R}^2 \rightarrow [0, 1] \mid n \in \mathbb{Z}, m \in \mathbb{Z}_+\}$$

on \mathbb{R}^2 by

$$\chi_{n,m} : \mathbb{R}^2 \rightarrow [0, 1], \quad \chi_{n,m}(\xi, \eta) = \rho_n(\eta) \cdot \chi_m(\theta_0^{-1} \cdot \langle n \rangle^{-2} \cdot \xi).$$

The support of the function $\chi_{n,m}$ is contained in the region

$$([-2^{m+1}\langle n \rangle^2\theta_0, -2^{m-1}\langle n \rangle^2\theta_0] \cup [2^{m-1}\langle n \rangle^2\theta_0, 2^{m+1}\langle n \rangle^2\theta_0]) \times I_n$$

when $m \geq 1$, and in $[-2\langle n \rangle^2\theta_0, 2\langle n \rangle^2\theta_0] \times I_n$ otherwise.

Definition 4.1. For a real number $r > 0$ and an integer $p \geq 1$, we define the norm $\|\cdot\|_{r,p}$ on the Schwartz space $\mathcal{S}(\mathbb{R}^2)$ by

$$(27) \quad \|u\|_{r,p} = \left(\sum_{n=-\infty}^{\infty} \sum_{m=0}^{\infty} (2^{rm} \cdot \|\mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F}u\|_{L^{2p}})^{2p} \right)^{1/2p}$$

where \mathcal{F} and $\mathcal{M}(\varphi)$ denote the Fourier transform and the multiplication operator by φ respectively, and $\|\cdot\|_{L^{2p}}$ denotes the L^{2p} norm. Let $\mathcal{B}^{r,p}(\mathbb{R}^2) \subset \mathcal{S}'(\mathbb{R}^2)$ be the completion of $\mathcal{S}(\mathbb{R}^2)$ with respect to this norm. For a subset $K \subset \mathbb{R}^2$, we write $\mathcal{B}^{r,p}(K)$ for the subspace of $\mathcal{B}^{r,p}(\mathbb{R}^2)$ that consists of elements whose support is contained in the closure of K .

Remark 4.2. We could introduce another parameter $q \in \mathbb{R}$ and define the Banach space $\mathcal{B}^{r,p,q}(\mathbb{R}^2)$ as the completion of $\mathcal{S}(\mathbb{R}^2)$ with respect to the norm

$$\|u\|_{r,p,q} = \left(\sum_{n=-\infty}^{\infty} \sum_{m=0}^{\infty} (2^{rm} \cdot \langle n \rangle^{2q} \cdot \|\mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F}u\|_{L^{2p}})^{2p} \right)^{1/2p}.$$

We can develop our argument presented below for these more general Banach spaces (regardless of the choice of q) in parallel, with slight differences in constants. One advantage of considering such generalization is that we can prove that the eigenfunctions of \mathcal{L}^t corresponding to the peripheral eigenvalues outside of the essential spectral radius belong to $C^\infty(X_f)$. This is essentially because $\cap_{r,q} \mathcal{B}^{r,p,q}(\mathbb{R}^2) = C^\infty(\mathbb{R}^2)$ and because the peripheral eigenvalues and the corresponding eigenfunctions do not depend on the choice of Banach spaces. But we restrict our argument below to the case $q = 0$ for simplicity's sake.

For technical argument in the next subsection, we introduce slight variants of the Banach space $\mathcal{B}^{r,p}(\mathbb{R}^2)$. For real numbers S and $E > 0$, let $A_{S,E} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear map defined by

$$(28) \quad A_{S,E} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Ex \\ SEx + y \end{pmatrix} = \begin{pmatrix} E & 0 \\ SE & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The transpose of its inverse is

$$A_{S,E}^\dagger \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} E^{-1} & -S \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

The Banach space $\mathcal{B}_{S,E}^{r,p}(\mathbb{R}^2)$ is defined as the push-forward of $\mathcal{B}^{r,p}(\mathbb{R}^2)$ by $A_{S,E}$. Precisely we define

$$\mathcal{B}^{r,p}(\mathbb{R}^2) = \{u \in \mathcal{D}'(\mathbb{R}^2) \mid u \circ A_{S,E} \in \mathcal{B}^{r,p}(\mathbb{R}^2)\}$$

and equip it with the norm

$$(29) \quad \|u\|_{r,p,S,E} := E^{1/2p} \cdot \|u \circ A_{S,E}\|_{r,p} \\ = \left(\sum_{n=-\infty}^{\infty} \sum_{m=0}^{\infty} (2^{rm} \|\mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m,S,E}) \circ \mathcal{F}u\|_{L^{2p}})^{2p} \right)^{1/2p}$$

where $\chi_{n,m,S,E} := \chi_{n,m} \circ (A_{S,E}^\dagger)^{-1}$.

4.2. Basic estimates. We provide a few basic lemmas related to the definitions introduced above. Note that the operator $\mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F}$ is written as a convolution operator

$$\mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F}u = \hat{\chi}_{n,m} * u$$

with $\hat{\chi}_{n,m} = (2\pi)^{-1} \mathcal{F}^{-1} \chi_{n,m}$.

Lemma 4.3. *For arbitrarily large $\nu > 0$, there exists a constant C_ν such that*

$$|\hat{\chi}_{n,m}(x, y)| \leq C_\nu \cdot (2^m \langle n \rangle^3) \cdot \langle 2^m \langle n \rangle^2 |x| \rangle^{-\nu} \cdot \langle \langle n \rangle \cdot |y| \rangle^{-\nu}$$

uniformly for integers n and $m \geq 0$. In particular, the L^1 norm of $\hat{\chi}_{n,m}$ is uniformly bounded.

Proof. The family of functions

$$X_{n,m}(\xi, \eta) := \chi_{n,m}(2^m \langle n \rangle^2 \xi, \langle n \rangle (\eta - n|n|))$$

for $n \in \mathbb{Z}$ and $m \in \mathbb{Z}_+$ are uniformly bounded in $\mathcal{S}(\mathbb{R}^2)$ and therefore so are the family of functions

$$\mathcal{F}^{-1} X_{n,m}(x, y) = (2^{-m} \langle n \rangle^{-3}) \cdot e^{in|n|y} \cdot \mathcal{F}^{-1} \chi_{n,m}(2^{-m} \langle n \rangle^{-2} x, \langle n \rangle^{-1} y).$$

This implies the conclusion of the lemma. \square

Similarly we have

Lemma 4.4. *The L^1 norm of $\hat{\chi}_{n,m,S,E} = (2\pi)^{-1} \mathcal{F}^{-1} \chi_{n,m,S,E}$ is bounded by a constant independent of n , m , S and E .*

By abuse of notation, we will write $\hat{\chi}_{n,m}$ also for the convolution operator by $\hat{\chi}_{n,m}$, so that $\hat{\chi}_{n,m}u = \hat{\chi}_{n,m} * u = \mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F}u$.

Lemma 4.5. *For integers n and $m \geq 0$ and for a bounded region $U \subset \mathbb{R}^2$, the convolution operator*

$$\hat{\chi}_{n,m} = \mathcal{F}^{-1} \circ \mathcal{M}(\chi_{n,m}) \circ \mathcal{F} : L^{2p}(U) \rightarrow L^{2p}(\mathbb{R}^2)$$

is a trace class operator. There exists a constant $C_0 > 0$, independent of n, m and U , such that

$$\|\hat{\chi}_{n,m} : L^{2p}(U) \rightarrow L^{2p}(\mathbb{R}^2)\|_{\text{Tr}} \leq C_0 \cdot 2^m \langle n \rangle^3 \cdot |U|_{n,m}$$

where $\|\cdot\|_{\text{Tr}}$ denotes the trace norm and

$$|U|_{n,m} := \int \max_{(x,y) \in U} (\langle 2^m \langle n \rangle^2 |x - x'| \rangle^{-2} \cdot \langle \langle n \rangle |y - y'| \rangle^{-2}) dx' dy'.$$

Remark 4.6. Since we consider operators between Banach spaces, it might be more standard to use the term “nuclear operator” and “nuclear norm” instead of “trace class operator” and “trace norm”. For the definition and basic properties of trace class (or nuclear) operators, we refer [4, Ch. 5].

Proof. Let us set $\chi'_{n,m} := \sum_{n',m'}^* \chi_{n',m'}$ where the sum \sum^* is taken over (n', m') such that $\text{supp } \chi_{n',m'} \cap \text{supp } \chi_{n,m} \neq \emptyset$. Since $\chi'_{n,m} \cdot \chi_{n,m} = \chi_{n,m}$, we may write the operator $\hat{\chi}_{n,m}$ as

$$\hat{\chi}_{n,m} u = \hat{\chi}'_{n,m} * \hat{\chi}_{n,m} * u = \int \phi_{z'} u dz'$$

where $\phi_{z'}$ is the rank one operator defined by

$$\phi_{z'} u(z) = \left(\int \hat{\chi}_{n,m}(z' - z'') u(z'') dz'' \right) \cdot \hat{\chi}'_{n,m}(z - z').$$

From Lemma 4.3, we have

$$\begin{aligned} \|\phi_{z'} : L^{2p}(U) \rightarrow L^{2p}(\mathbb{R}^2)\|_{\text{Tr}} \\ \leq C_0 \langle n \rangle^3 2^m \cdot \max_{(x,y) \in U} (\langle 2^m \langle n \rangle^2 |x - x'| \rangle^{-2} \langle \langle n \rangle |y - y'| \rangle^{-2}) \end{aligned}$$

for $z' = (x', y')$. Indeed the left hand side is bounded by

$$\|\hat{\chi}_{n,m}\|_{L^{2p}} \cdot \|\hat{\chi}'_{n,m}\|_{(L^{2p})^*} = \|\chi_{n,m}\|_{L^{2p}} \cdot \|\chi'_{n,m}\|_{L^q}$$

with $q > 0$ such that $q^{-1} + (2p)^{-1} = 1$ and hence by $C_0 \langle n \rangle^3 2^m$ at least. Because $\nu > 0$ in Lemma 4.3 is arbitrarily large, we can get the latter term on the right hand side in addition. Finally we obtain the lemma by the triangle inequality. \square

For the purpose of extracting the low-frequency part of functions, we consider the operators

$$\mathcal{K}_k : \mathcal{S}'(\mathbb{R}^2) \rightarrow \mathcal{S}(\mathbb{R}^2), \quad \mathcal{K}_k u = \sum_{n,m: 2^m \langle n \rangle^2 \leq k} \hat{\chi}_{n,m} u$$

for integers $k > 0$. If $U \subset \mathbb{R}^2$ is a bounded region, the operator $\mathcal{K}_k : \mathcal{B}^{r,p}(U) \rightarrow \mathcal{B}^{r,p}(\mathbb{R}^2)$ is a trace class operator from Lemma 4.5 and hence compact.

As a model of the semiflow T_f^t viewed in local charts (that we will choose in the next section), we consider a C^∞ diffeomorphism

$$(30) \quad A : V \rightarrow A(V) \subset \mathbb{R}^2, \quad A(x, y) = (Ex, y + g(Ex))$$

where $E \geq 1$, $V := (-E^{-1}\eta_*, E^{-1}\eta_*) \times \mathbb{R} \subset \mathbb{R}^2$ with some small $\eta_* > 0$ and $g : (-\eta_*, \eta_*) \rightarrow \mathbb{R}$ is a C^∞ function satisfying $|g'(x)| \leq \gamma_0 \theta_0$. Letting $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ be

a C^∞ function whose support is contained in $(-\eta_*, \eta_*) \times \mathbb{R}$, we consider the transfer operator

$$(31) \quad L : C_0^\infty(V) \rightarrow C_0^\infty(A(V)), \quad Lu = (\varphi \cdot u) \circ A^{-1}.$$

In the next proposition, we suppose that the function $\varphi(x, y) \in C_0^\infty(\mathbb{R}^2)$ satisfies

$$(32) \quad \left\| \frac{\partial^m \varphi}{\partial y^m} \right\|_\infty \leq K_m \quad \text{for } m \geq 0$$

for some given constants $K_m > 0$. When we apply the proposition below in the next section, we will consider many different functions as φ , which satisfy the condition (32) for some uniform constants K_m .

Proposition 4.7. *If we have (in addition to the setting above) that*

$$(33) \quad |g'(x) - g'(0)| < (1 - \gamma_0)\theta_0/E \quad \text{for all } x \in (-\eta_*, \eta_*),$$

the operator L extends to a bounded operator

$$L : \mathcal{B}^{r,p}(V) \rightarrow \mathcal{B}_{S,E}^{r,p}(A(\text{supp } \varphi)) \quad \text{where } S = g'(0).$$

There exists a constant $C_0 > 0$, which depends only on p, r and the constants K_m 's in (32), such that we have

$$\|L \circ (1 - \mathcal{K}_k) : \mathcal{B}^{r,p}(V) \rightarrow \mathcal{B}_{S,E}^{r,p}(A(\text{supp } \varphi))\| \leq C_0 E^{1/2p}$$

provided that we take sufficiently large $k > 0$ according to A and φ .

Proof. Recall the linear map $A_{S,E}$ in (28). The diffeomorphism $A_{S,E}^{-1} \circ A$ satisfies the assumption on A for the case $E = 1$ and $S = 0$. Recall also that $\mathcal{B}_{S,E}^{r,p}(\mathbb{R}^2)$ is defined as the push-forward of $\mathcal{B}^{r,p}(\mathbb{R}^2)$ by $A_{S,E}$ and equipped with the norm (29) having the factor $E^{1/2p}$. Hence, to prove the statement of the lemma, it is enough to prove it in the case $E = 1$ and $S = 0$ (and $\mathcal{B}_{S,E}^{r,p}(\mathbb{R}^2) = \mathcal{B}^{r,p}(\mathbb{R}^2)$ consequently).

We assume $E = 1$ and $S = 0$. Take $u \in \mathcal{S}(\mathbb{R}^2)$ arbitrarily and set

$$\begin{aligned} u_{n,m} &= \hat{\chi}_{n,m}(u), \\ v_{(n,m) \rightarrow (n',m')} &= \hat{\chi}_{n',m'}(Lu_{n,m}) = \hat{\chi}_{n',m'} \circ L \circ \hat{\chi}'_{n,m}(u_{n,m}) \end{aligned}$$

and

$$v_{n',m'} = \hat{\chi}_{n',m'}(Lu) = \sum_{(n,m)} v_{(n,m) \rightarrow (n',m')}$$

where $\hat{\chi}'_{n,m}$ is that defined in the proof of Lemma 4.5 and $\hat{\chi}'_{n,m}$ denotes the convolution operator by the function $(2\pi)^{-1} \mathcal{F}^{-1} \chi'_{n,m}$. Since $(1 - \mathcal{K}_k)$ on $\mathcal{B}^{r,p}(\mathbb{R}^2)$ is bounded uniformly in k and cut off the low-frequency components, it suffices to show

$$(34) \quad \sum_{n',m'} (2^{rm'} \|v_{n',m'}\|_{L^{2p}})^{2p} \leq C_0 \sum_{n,m} (2^{rm} \|u_{n,m}\|_{L^{2p}})^{2p}$$

assuming that $u_{n,m}$ vanishes when $2^m \langle n \rangle^2 \leq k$ for some large k .

We estimate the operator norm of $\hat{\chi}_{n',m'} \circ L \circ \hat{\chi}'_{n,m}$ on $L^{2p}(\mathbb{R}^2)$. Let us set

$$(35) \quad \Delta_1(n, n') = \begin{cases} 1, & \text{if } |n - n'| \leq 3; \\ \max\{n, n'\}, & \text{otherwise} \end{cases}$$

and

$$\Delta_2(n, m, n', m') = \begin{cases} 1, & \text{if either } m' = 0 \text{ or } 2^{m'} \langle n' \rangle^2 \leq 2^{m+4} \langle n \rangle^2; \\ \max\{2^m \langle n \rangle^2, 2^{m'} \langle n' \rangle^2\}, & \text{otherwise.} \end{cases}$$

We are going to prove two estimates: One is that, for any $\nu > 0$, there exists a constant $C_\nu > 0$, depending only on ν and the constants K_m 's in (32), such that

$$(36) \quad \|\hat{\chi}_{n', m'} \circ L \circ \hat{\chi}'_{n, m}\|_{L^{2p}} \leq C_\nu \Delta_1(n, n')^{-\nu}$$

for any combination (n, m, n', m') . The other is that, for any $\nu > 0$, there exists a constant $C(A, \varphi, \nu)$, depending ν , A and φ , such that

$$(37) \quad \|\hat{\chi}_{n', m'} \circ L \circ \hat{\chi}'_{n, m}\|_{L^{2p}} \leq C(A, \varphi, \nu) \cdot \Delta_1(n, n')^{-\nu} \cdot \Delta_2(n, m, n', m')^{-\nu}$$

for any combination (n, m, n', m') .

The required estimate (34) will follow from (36) and (37). By using Hölder inequality, we see that the left hand side of (34) is bounded by

$$\sum_{n', m'} 2^{2pr m'} \left\| \sum_{n, m} v_{(n, m) \rightarrow (n', m')} \right\|_{L^{2p}}^{2p} \leq \sum_{n', m'} \sum_{n, m} \Delta_{n, m, n', m'} \cdot 2^{2pr m'} \|v_{(n, m) \rightarrow (n', m')}\|_{L^{2p}}^{2p}$$

where

$$\Delta_{n, m, n', m'} = C_0 \cdot 2^{2pr(m' - m)} \cdot 2^{|m' - m|} \cdot \langle n - n' \rangle^2$$

with $C_0 > 0$ a constant depending only on p and r . The estimate (37) with large ν implies that the components $v_{(n, m) \rightarrow (n', m')}$ is very small if $m' > 0$ and $2^{m'} \langle n' \rangle^2 > 2^{m+4} \langle n \rangle^2$. (Recall that we suppose $u_{n, m}$ vanishes when $2^m \langle n \rangle^2 \leq k$ for some large k .) Hence the sum on the right hand side above over such combinations (n, m, m', m') are negligible or more precisely bounded by $c \cdot \sum_{n, m} (2^{rm} \|u_{n, m}\|_{L^{2p}})^{2p}$ and we may let the constant $c > 0$ be arbitrarily small by letting k large. To the remaining components for which either $m' = 0 \leq m$ or $2^{m'} \langle n' \rangle^2 \leq 2^{m+4} \langle n \rangle^2$ holds, we apply the estimate (36) with large ν . Then we obtain the required estimate (34) by elementary computation.

To prove (36) and (37), we look into the integral kernel of $\hat{\chi}_{n', m'} \circ L \circ \hat{\chi}'_{n, m}$ and estimate it by using integration by parts. Though the following argument is elementary and already presented in [1], we give it to some detail for completeness. (We will use a similar argument later, where we will omit the proof.) To begin with, let us make the following observation which motivates the definitions of $\Delta_1(\cdot)$ and $\Delta_2(\cdot)$: There exists a small constant $c > 0$ such that, for any $(\xi', \eta') \in \text{supp } \chi_{n', m'}$ and any $(\tilde{\xi}, \tilde{\eta}) \in DA_w^\dagger(\text{supp } \chi'_{n, m})$ with $w \in V$, we have

$$(38) \quad |\eta' - \tilde{\eta}| \geq c \max\{|n|, |n'|\} \quad \text{if} \quad |n - n'| \geq 4$$

and

$$(39) \quad |\xi' - \tilde{\xi}| \geq c \max\{2^m \langle n \rangle^2, 2^{m'} \langle n' \rangle^2\} \quad \text{if} \quad m' > 0 \quad \text{and} \quad 2^{m'} \langle n' \rangle^2 > 2^{m+4} \langle n \rangle^2.$$

Next let us write the operator $\hat{\chi}_{n', m'} \circ L \circ \hat{\chi}'_{n, m}$ as an integral operator

$$\hat{\chi}_{n', m'} \circ L \circ \hat{\chi}'_{n, m} u(z') = (2\pi)^{-2} \int K(z', z) u(z) dz$$

with the integral kernel

$$(40) \quad K(z', z) = \int e^{i\theta' \cdot (z' - w) + i\theta \cdot (A^{-1}(w) - z)} \cdot \chi_{n', m'}(\theta') \cdot \chi'_{n, m}(\theta) \cdot \varphi(A^{-1}(w)) d\theta d\theta' dw.$$

To apply integration by parts, we consider the differential operators

$$\mathcal{D}_1 = \frac{1 - i(\eta - \eta') \cdot \partial_y}{1 + |\eta - \eta'|^2}, \quad \mathcal{D}_2 = \frac{1 - i(DA_w^\dagger \theta - \theta') \cdot \partial_w}{1 + |DA_w^\dagger \theta - \theta'|^2}$$

expressed in the coordinates $\theta = (\xi, \eta)$, $\theta' = (\xi', \eta')$ and $w = (x, y)$. These satisfy

$$\mathcal{D}_j e^{i(\theta \cdot A^{-1}(w) - \theta' \cdot w)} = e^{i(\theta \cdot A^{-1}(w) - \theta' \cdot w)}, \quad j = 1, 2.$$

(For the case $j = 1$, note that A is written in the form (30).) Hence

$$\begin{aligned} \int e^{i(\theta' \cdot A^{-1}(w) - \theta \cdot w)} \Phi(w) dw &= \int \left(\mathcal{D}_j e^{i(\theta' \cdot A^{-1}(w) - \theta \cdot w)} \right) \Phi(w) dw \\ &= \int e^{i(\theta' \cdot A^{-1}(w) - \theta \cdot w)} \cdot ({}^t \mathcal{D}_j) \Phi(w) dw \end{aligned}$$

for $j = 1, 2$, where ${}^t \mathcal{D}_j$ denotes the transpose of \mathcal{D}_j with respect to the L^2 inner product. We apply this formula with $j = 1$ for several times if $|n - n'| \geq 4$ and then apply that with $j = 2$ for several times if $m' > 0$ and $2^{m'} \langle n' \rangle^2 > 2^{m+4} \langle n \rangle^2$. As the result, we get the expression of the form

$$K(z', z) = \int e^{i\theta' \cdot (z' - w) + i\theta \cdot (A^{-1}(w) - z)} \cdot \Psi(w, \theta, \theta') dw d\theta d\theta'$$

where the integration with respect to the variables θ' and θ are taken over the supports of $\chi_{n', m'}$ and $\chi'_{n, m}$ respectively. Using the estimates (38) and (39), we see, for arbitrarily large $\nu \geq 1$ and for any integers $\alpha, \alpha', \beta, \beta' \geq 0$, that

$$|\partial_\xi^\alpha \partial_\eta^\beta \partial_{\xi'}^{\alpha'} \partial_{\eta'}^{\beta'} \Psi(w, \theta, \theta')| \leq \frac{C_{\nu, \alpha, \beta, \alpha', \beta'} \cdot \Delta_1(n, n')^{-\nu} \cdot \Delta_2(n, m, n', m')^{-\nu}}{\langle n \rangle^\beta \cdot \langle n' \rangle^{\beta'} \cdot \langle 2^m \langle n \rangle^2 \rangle^\alpha \cdot \langle 2^{m'} \langle (n')^2 \rangle \rangle^{\alpha'}}$$

where the constants $C_{\nu, \alpha, \beta, \alpha', \beta'}$ depend on A and φ but not on n, m, n' nor m' . This implies that, for arbitrarily large $\nu > 0$, we have

$$(41) \quad |K(z', z)| \leq C(A, \varphi, \nu) \cdot \Delta_1(n, n')^{-\nu} \cdot \Delta_2(n, m, n', m')^{-\nu} \cdot \int \rho_{n', m'}^{(\nu)}(z' - w) \cdot \rho_{n, m}^{(\nu)}(A^{-1}(w) - z) dw$$

where

$$\rho_{n, m}^{(\nu)}(x, y) = 2^m \langle n \rangle^3 \cdot \langle 2^m \langle n \rangle^2 |x - x'| \rangle^{-\nu} \cdot \langle \langle n \rangle |y - y'| \rangle^{-\nu}.$$

Hence we conclude the estimate (37) by Young's inequality. Note that if we did not apply integration by parts using \mathcal{D}_2 , we obtain the estimate

$$(42) \quad |\partial_\xi^\alpha \partial_\eta^\beta \partial_{\xi'}^{\alpha'} \partial_{\eta'}^{\beta'} \Psi(w, \theta, \theta')| \leq \frac{C'_{\nu, \alpha, \beta, \alpha', \beta'} \cdot \Delta_1(n, n')^{-\nu}}{\langle n \rangle^\beta \cdot \langle n' \rangle^{\beta'} \cdot \langle 2^m \langle n \rangle^2 \rangle^\alpha \cdot \langle 2^{m'} \langle (n')^2 \rangle \rangle^{\alpha'}}$$

where the constants $C'_{\nu, \alpha, \beta, \alpha', \beta'}$ depend on ν and the constants K_m 's in (32) but not on A, φ, n, m, n' nor m' . Hence we obtain (36) by a parallel argument. \square

Lemma 4.8. *Let $U \subset \mathbb{R}^2$ be a bounded region. Let $\rho_i : \mathbb{R}^2 \rightarrow [0, 1]$, $1 \leq i \leq I$, be a finite set of C^∞ functions with compact supports such that $\sum_{i=1}^I \rho_i(x) \equiv 1$ for $x \in U$. Suppose that the functions ρ_i satisfy the condition (32) with $\varphi = \rho_i$ for some constants K_m for $m \geq 0$. Then there exists a constant $C_0 > 0$, which depends only on p, r and the constants K_m , such that, for sufficiently large $k > 0$ (depending on the functions ρ_j), we have*

$$\sum_{i=1}^I \|\rho_i \cdot (1 - \mathcal{K}_k)u\|_{r,p}^{2p} \leq C_0 \|u\|_{r,p}^{2p}$$

and

$$\|(1 - \mathcal{K}_k)u\|_{r,p}^{2p} \leq C_0 \mu^{2p-1} \cdot \sum_{i=1}^I \|\rho_i \cdot u\|_{r,p}^{2p}$$

for any $u \in \mathcal{B}^{r,p}(U)$, where μ is the intersection multiplicity of the subsets

$$X_i := \{x \in \mathbb{R} \mid (x, y) \in \text{supp } \rho_i \text{ for some } y \in \mathbb{R}\} \quad \text{for } 1 \leq i \leq I.$$

Proof. To get the claims of the lemma, we reconsider the argument in the proof of Proposition 4.7 in the case $A = \text{Id}$ and $\varphi = \rho_i$, and pay extra attention to the localized property of the kernel of $\hat{\chi}_{n',m'} \circ L \circ \hat{\chi}'_{n,m}$ given in (41) and (42). We omit the detail of the proof as it is easy to provide. \square

4.3. An L^p estimate using transversality. The next lemma is the key step of the argument in the proof of Theorem 3.5.

Proposition 4.9. *Let $S(i)$ and $E(i)$, $1 \leq i \leq M$, be real numbers such that $|S(i)| \leq \gamma_0 \theta_0$ and $E(i) \geq \ell$. For a p -tuple $\mathbf{i} = (\mathbf{i}(1), \mathbf{i}(2), \dots, \mathbf{i}(p)) \in \{1, 2, \dots, M\}^p$, we define*

$$S(\mathbf{i}) := \sum_{k=1}^p S(\mathbf{i}(k)), \quad E(\mathbf{i}) := \left(\sum_{k=1}^p E(\mathbf{i}(k))^{-1} \right)^{-1}$$

and set

$$(43) \quad \Delta = \max_{\xi} \sum_{\mathbf{i} \in \{1, 2, \dots, M\}^p} \langle (E(\mathbf{i})/\theta_0) |\xi - S(\mathbf{i})| \rangle^{-r}.$$

Then there exists a constant $C_0 > 0$, independent of $S(i)$ and $E(i)$, such that, for sufficiently large $k > 0$, we have

$$(44) \quad \left\| \sum_{i=1}^M (1 - \mathcal{K}_k)u_i \right\|_{r,p}^{2p} \leq C_0 \max \left\{ \frac{M^{2p-1}}{\min_{1 \leq i \leq M} E(i)^{2pr}}, M^{p-1} \Delta \right\} \sum_{i=1}^M \|u_i\|_{r,p,S(i),E(i)}^{2p}$$

for any $u_i \in \mathcal{B}_{S(i),E(i)}^{r,p}(\mathbb{R}^2)$.

Proof. Inspecting the supports of the functions $\chi_{n,m}$ and $\chi_{n',m',S(i),E(i)}$, we find a constant $c_0 > 0$, independent of $S(i)$ and $E(i)$, such that

$$\chi_{n,m} \cdot \chi_{n',m',S(i),E(i)} \equiv 0 \quad (\text{or } \hat{\chi}_{n,m} * \hat{\chi}_{n',m',S(i),E(i)} = 0)$$

if $|n - n'| \geq 3$ or if $m > 0$ and $m \geq m' - \log E(i)/\log 2 + c_0$. From Lemma 4.4, the L^1 norm of the functions $\hat{\chi}_{n,m}$ and $\hat{\chi}_{S(i),E(i),n',m'}$ are bounded by a constant independent of, $n, m, n', m', S(i)$ and $E(i)$ and therefore so are the operator

norms of the convolution operators with these functions on $L^{2p}(\mathbb{R}^2)$. Hence, by using Hölder inequality twice, we obtain that

$$\begin{aligned}
 (45) \quad & \sum_n \sum_{m>0} \left(2^{rm} \left\| \hat{\chi}_{n,m} \left(\sum_{i=1}^M u_i \right) \right\|_{L^{2p}} \right)^{2p} \\
 & \leq M^{2p-1} \sum_n \sum_{m>0} \sum_{i=1}^M 2^{2prm} \|\hat{\chi}_{n,m}(u_i)\|_{L^{2p}}^{2p} \\
 & \leq M^{2p-1} \sum_{i=1}^M \sum_n \sum_{m>0} \left\| \sum_{n'} \sum_{m'} 2^{rm} \hat{\chi}_{n,m} \circ \hat{\chi}_{n',m',S(i),E(i)}(u_i) \right\|_{L^{2p}}^{2p} \\
 & \leq \frac{C_0 \cdot M^{2p-1}}{\min_{1 \leq i \leq M} E(i)^{2pr}} \sum_{i=1}^M \sum_{n'} \sum_{m'} 2^{2prm'} \|\hat{\chi}_{n',m',S(i),E(i)}(u_i)\|_{L^{2p}}^{2p}.
 \end{aligned}$$

Notice that we excluded the components with $m = 0$ in the estimate above. Below we give an estimate on the components with $m = 0$, which is more essential. Note that we may (and will) suppose that $|n|$ is large, by letting k (in the definition of \mathcal{K}_k) be larger if necessary.

For a p -tuple $\mathbf{i} = (\mathbf{i}(1), \mathbf{i}(2), \dots, \mathbf{i}(p)) \in \{1, 2, \dots, M\}^p$, we write

$$u_{\mathbf{i}} = \prod_{k=1}^p \hat{\chi}_{n,0}(u_{\mathbf{i}(k)})$$

and estimate the L^2 norm of $\hat{\chi}_{S(\mathbf{i}),E(\mathbf{i}),\tilde{n},\tilde{m}} * u_{\mathbf{i}}$ for integers \tilde{n} and $\tilde{m} \geq 0$. Since the support of $\mathcal{F}u_{\mathbf{i}}$ is contained in the subset

$$\sum_{k=1}^p \text{supp } \chi_{n,0} := \left\{ \sum_{k=1}^p x_i \mid x_i \in \text{supp } \chi_{n,0} \right\} \subset \mathbb{R}^2,$$

we have $\hat{\chi}_{S(\mathbf{i}),E(\mathbf{i}),\tilde{n},\tilde{m}} * u_{\mathbf{i}} = 0$ unless

$$(46) \quad ||\tilde{n}|^2 - p|n|^2| \leq p(2|n| + 1) + 2|\tilde{n}| + 1.$$

We henceforth suppose that \tilde{n} satisfies (46). Since we assume $|n|$ is large, this implies that the ratio \tilde{n}/n is close to \sqrt{p} and we have $|\tilde{n} - \sqrt{p}n| \leq 3(\sqrt{p} + 1)$.

For convenience in the argument below, we introduce the functions

$$\zeta_{n,m,S,E}(\xi, \eta) := \chi_{n,m,S,E}(\xi, \eta) \cdot \chi(\theta_0^{-1} \cdot \langle n \rangle^{-2} \cdot \xi)$$

on \mathbb{R}^2 , which satisfy

$$\sum_{m \geq 0} \zeta_{n,m,S,E}(\xi, \eta) = \rho_n(\eta) \cdot \chi(\theta_0^{-1} \cdot \langle n \rangle^{-2} \cdot \xi) = \chi_{n,0}(\xi, \eta).$$

We write $\hat{\zeta}_{n,m,S,E}$ for the convolution operator by the function $(2\pi)^{-1} \mathcal{F}^{-1} \zeta_{n,m,S,E}$. Then we have

$$(47) \quad \sum_{m \geq 0} \hat{\zeta}_{n,m,S,E}(u) = \hat{\chi}_{n,0}(u)$$

and also

$$\|\hat{\zeta}_{n,m,S,E}(u)\|_{L^{2p}} \leq C_0 \|\hat{\chi}_{n,m,S,E}(u)\|_{L^{2p}}$$

for a constant $C_0 > 0$ which depend only on the choice of the function $\chi(\cdot)$.

For a sequence $\mathbf{m} = (\mathbf{m}(1), \dots, \mathbf{m}(p)) \in (\mathbb{Z}_{\geq 0})^p$ of non-negative integers, put

$$|\mathbf{m}| = \max_{1 \leq i \leq p} \mathbf{m}(i).$$

By inspecting the position of the supports of $\chi_{n,m,S,E}(\cdot)$ in the ξ -coordinate, we find a constant $C > 0$, which depends only on p , such that, if $|\mathbf{m}| < \tilde{m} - C$, we have

$$\text{supp } \hat{\chi}_{\tilde{n}, \tilde{m}, S(\mathbf{i}), E(\mathbf{i})} \cap \left(\sum_{k=1}^p \text{supp } \hat{\chi}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))} \right) = \emptyset$$

and hence

$$\hat{\chi}_{\tilde{n}, \tilde{m}, S(\mathbf{i}), E(\mathbf{i})} * \left(\prod_{k=1}^p \hat{\zeta}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)}) \right) = 0.$$

From this and (47), we have

$$\|\hat{\chi}_{\tilde{n}, \tilde{m}, S(\mathbf{i}), E(\mathbf{i})}(u_{\mathbf{i}})\|_{L^2}^2 = \left\| \sum_{|\mathbf{m}| \geq \tilde{m} - C} \prod_{k=1}^p \hat{\zeta}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)}) \right\|_{L^2}^2.$$

By using Schwarz and Hölder inequality, we continue

$$\begin{aligned} &\leq C_0 \sum_{|\mathbf{m}| \geq \tilde{m} - C} 2^{|\mathbf{m}|} \left\| \prod_{k=1}^p \hat{\zeta}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)}) \right\|_{L^2}^2 \\ &\leq C_0 \sum_{|\mathbf{m}| \geq \tilde{m} - C} 2^{|\mathbf{m}|} \prod_{k=1}^p \left\| \hat{\zeta}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)}) \right\|_{L^{2p}}^2 \end{aligned}$$

and further

$$\begin{aligned} &\leq C_0 2^{-r\tilde{m}} \sum_{\mathbf{m}} \prod_{k=1}^p \left(2^{(r+1)\mathbf{m}(k)} \|\hat{\zeta}_{n, \mathbf{m}(k), S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)})\|_{L^{2p}}^2 \right) \\ &\leq C_0 2^{-r\tilde{m}} \prod_{k=1}^p \left(\sum_{m=0}^{\infty} 2^{(r+1)m} \|\hat{\chi}_{n, m, S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)})\|_{L^{2p}}^2 \right) \end{aligned}$$

where (and henceforth in the proof below) $C_0 > 0$ denotes constants depending only on p and r and its values may be different from place to place. We therefore conclude

$$\begin{aligned} (48) \quad &2^{r\tilde{m}} \|\hat{\chi}_{\tilde{n}, \tilde{m}, S(\mathbf{i}), E(\mathbf{i})}(u_{\mathbf{i}})\|_{L^2}^2 \\ &\leq C_0 \prod_{k=1}^p \left(\sum_{m=0}^{\infty} 2^{2rm} \|\hat{\chi}_{n, m, S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)})\|_{L^{2p}}^2 \right) \end{aligned}$$

where we used the fact $2r \geq r+1$ that follows from (16).

Now we are going to prove the conclusion of the proposition. Recall the quantity Δ defined in (43) and write

$$W_{\mathbf{i}}(\xi, \eta) = \langle (E(\mathbf{i})/\theta_0) | \xi / \langle \eta \rangle - S(\mathbf{i}) | \rangle^{r/2}.$$

Then we have

$$\begin{aligned}
\left\| \hat{\chi}_{n,0} \left(\sum_{i=1}^M u_i \right) \right\|_{L^{2p}}^{2p} &= \left\| \sum_{\mathbf{i}} u_{\mathbf{i}} \right\|_{L^2}^2 \leq \sum_{\mathbf{i}, \mathbf{j}} \|W_{\mathbf{j}}^{-1} \cdot W_{\mathbf{i}} \cdot \mathcal{F}u_{\mathbf{i}}\|_{L^2} \cdot \|W_{\mathbf{i}}^{-1} \cdot W_{\mathbf{j}} \cdot \mathcal{F}u_{\mathbf{j}}\|_{L^2} \\
&\leq \sum_{\mathbf{i}, \mathbf{j}} \|W_{\mathbf{j}}^{-1} \cdot W_{\mathbf{i}} \cdot \mathcal{F}u_{\mathbf{i}}\|_{L^2}^2 \\
&\leq \left\| \sum_{\mathbf{j}} W_{\mathbf{j}}^{-2} \right\|_{\infty} \cdot \sum_{\mathbf{i}} \|W_{\mathbf{i}} \cdot \mathcal{F}u_{\mathbf{i}}\|_{L^2}^2 \leq \Delta \cdot \sum_{\mathbf{i}} \|W_{\mathbf{i}} \cdot \mathcal{F}u_{\mathbf{i}}\|_{L^2}^2.
\end{aligned}$$

Since $W_{\mathbf{i}}(\xi, \eta) \leq C_0 2^{r\tilde{m}/2}$ on the support of $\chi_{\tilde{n}, \tilde{m}, S(\mathbf{i}(k)), E(\mathbf{i}(k))}$, we have from (48) that

$$\begin{aligned}
\|W_{\mathbf{i}} \cdot \mathcal{F}u_{\mathbf{i}}\|_{L^2}^2 &\leq C_0 \sum_{\tilde{n}: |\tilde{n}-n| \leq 3(\sqrt{p}+1)} \sum_{\tilde{m}=0}^{\infty} 2^{r\tilde{m}} \|\hat{\chi}_{\tilde{n}, \tilde{m}, S(\mathbf{i}), E(\mathbf{i})}(u_{\mathbf{i}})\|_{L^2}^2 \\
&\leq C_0 \sum_{\tilde{n}: |\tilde{n}-n| \leq 3(\sqrt{p}+1)} \prod_{k=1}^p \left(\sum_{m=0}^{\infty} 2^{2rm} \|\hat{\chi}_{n, m, S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)})\|_{L^{2p}}^2 \right).
\end{aligned}$$

From the last two inequalities, we deduce

$$\begin{aligned}
(49) \quad &\sum_n \left\| \hat{\chi}_{n,0} \left(\sum_{i=1}^M u_i \right) \right\|_{L^{2p}}^{2p} \\
&\leq C_0 \Delta \cdot \sum_n \sum_{\tilde{n}: |\tilde{n}-n| \leq 3(\sqrt{p}+1)} \sum_{\mathbf{i}} \prod_{k=1}^p \left(\sum_{m=0}^{\infty} 2^{2rm} \|\hat{\chi}_{n, m, S(\mathbf{i}(k)), E(\mathbf{i}(k))}(u_{\mathbf{i}(k)})\|_{L^{2p}}^2 \right) \\
&\leq C_0 \Delta \cdot \sum_n \left(\sum_{i=1}^M \sum_{m=0}^{\infty} 2^{2rm} \|\hat{\chi}_{n, m, S(\mathbf{i}), E(\mathbf{i})}(u_{\mathbf{i}})\|_{L^{2p}}^2 \right)^p \\
&\leq C_0 \Delta M^{p-1} \cdot \sum_{i=1}^M \sum_n \sum_{m=0}^{\infty} 2^{2prm} \|\hat{\chi}_{n, m, S(\mathbf{i}), E(\mathbf{i})}(u_{\mathbf{i}})\|_{L^{2p}}^{2p} \quad \text{by Hölder inequality.}
\end{aligned}$$

Finally note that

$$\left\| \sum_{i=1}^M (1 - \mathcal{K}_k) u_i \right\|_{r,p}^{2p} \leq \sum_{n,m}^* \left(2^{rm} \left\| \hat{\chi}_{n,m} \left(\sum_{i=1}^M u_i \right) \right\|_{L^{2p}} \right)^{2p}$$

where the sum $\sum_{n,m}^*$ is taken over n and $m \geq 0$ such that $2^m \langle n^2 \rangle \geq k$. By (45) and (49), we obtain the conclusion of the proposition. \square

Corollary 4.10. *In Proposition 4.7, the operator L is a bounded operator from $\mathcal{B}^{r,p}(V)$ to $\mathcal{B}^{r,p}(A(\text{supp } \varphi))$.*

Proof. It suffices to show that the inclusion $\iota : \mathcal{B}_{S,E}^{r,p}(A(\text{supp } \varphi)) \rightarrow \mathcal{B}^{r,p}(A(\text{supp } \varphi))$ is bounded. This follows from Proposition 4.9 applied to the trivial case $M = 1$. \square

5. PROOF OF THEOREM 3.5

Below we set up a system of local charts on X_f so that the flow T_f^t looks smooth in each of them and then introduce the Banach space $\mathcal{B}^{r,p}(X_f)$ using such local charts and the Banach space $\mathcal{B}^{r,p}(\mathbb{R}^2)$. Once we have done with these, the proof of Theorem 3.5 is not very difficult and obtained basically by applying the propositions (especially Proposition 4.9) to the transfer operators induced on the local charts. Unfortunately a slight combinatorial complication is caused by the fact that we admit the “exceptional set” \mathcal{E} in the definition of $\mathcal{G}(J, n, \varepsilon, \delta; p)$. In order to present the idea of the proof clearly, we first prove the conclusion of the theorem assuming a stronger condition where $\mathcal{E} = \emptyset$ in Subsection 5.4 and then explain how we modify the argument to obtain the theorem in Subsection 5.5.

5.1. System of local charts on X_f and the definition of $\mathcal{B}^{r,p}(X_f)$. To begin with, we take two small real numbers $\eta_0 > 0$ and $\delta_0 > 0$ and consider the open rectangle

$$R = (-\eta_0, \eta_0) \times (4\delta_0, 7\delta_0) \subset Q = (-3\eta_0, 3\eta_0) \times (0, 11\delta_0).$$

For each $a = (x_0, y_0) \in X_f$, we consider the two mappings

$$\tilde{\kappa}_a : Q \rightarrow S^1 \times \mathbb{R}, \quad \tilde{\kappa}_a(x, y) = (x_0 + x, y_0 + y).$$

and

$$\kappa_a := \pi \circ \tilde{\kappa}_a : Q \rightarrow X_f$$

where

$$(50) \quad \pi : S^1 \times \mathbb{R}_+ \rightarrow X_f, \quad \pi(x, y) = (\tau^{n(x,y;f)}(x), y - f^{(n(x,y;f))})$$

and $\mathbb{R}_+ = \{s \in \mathbb{R} \mid s \geq 0\}$. (See Figure 2.) We suppose that η_0 and δ_0 are so small that both of κ_a and $\tilde{\kappa}_a$ are injective for any $a \in X_f$.

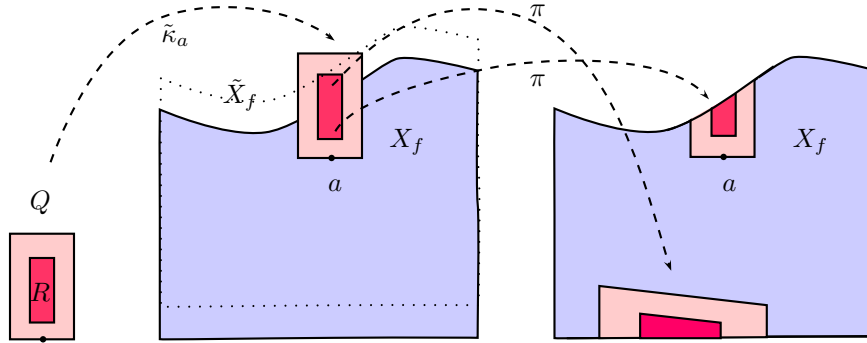


FIGURE 2. The mappings $\tilde{\kappa}_a$, π and κ_a .

Next we take a finite subset A of X_f so that the images $\tilde{\kappa}_a(R)$ for $a \in A$ cover the subset

$$\tilde{X}_f := \{(x, y) \in S^1 \times \mathbb{R}_+ \mid 5\delta_0 \leq y \leq f(x) + 6\delta_0\}.$$

Letting δ_0 and the ratio η_0/δ_0 be small, we may and do assume that the intersection multiplicity of $\{\tilde{\kappa}_a(R)\}_{a \in A}$ is bounded by an absolute constant (say, by 4).

We define the Banach space $\mathcal{B}^{r,p}(X_f)$ as follows. We suppose that the product space $\bigoplus_{a \in A} \mathcal{B}^{r,p}(R)$ is a Banach space with the norm

$$\|\mathbf{u}\|_{r,p} = \left(\sum_{a \in A} \|u_a\|_{r,p}^{2p} \right)^{1/2p} \quad \text{for } \mathbf{u} = (u_a)_{a \in A} \in \bigoplus_{a \in A} \mathcal{B}^{r,p}(R).$$

Then the operator

$$(51) \quad \Pi : \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \rightarrow L^2(X_f), \quad \Pi((\varphi_a)_{a \in A}) = \sum_{a \in A} \varphi_a \circ \kappa_a^{-1}$$

is bounded because $\mathcal{B}^{r,p}(R) \subset \mathcal{B}^{r,2}(R) \subset L^2(R)$.

Definition 5.1. Let $\mathcal{B}^{r,p}(X_f) \subset L^2(X_f)$ be the image of (51). This is a Banach space with respect to the norm

$$\|u\|_{\mathcal{B}^{r,p}} = \inf \left\{ \|\mathbf{u}\|_{r,p} \mid u = \Pi(\mathbf{u}), \mathbf{u} \in \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \right\}.$$

The operator Π in (51) is then restricted to a bounded operator

$$\Pi : \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \rightarrow \mathcal{B}^{r,p}(X_f)$$

with operator norm 1.

We next define a bounded operator $\mathbf{I} : \mathcal{B}^{r,p}(X_f) \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R)$ which makes the following diagram with $t = 6\delta_0$ commutes:

$$(52) \quad \begin{array}{ccc} & \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) & \\ \mathbf{I} \nearrow & & \downarrow \Pi \\ \mathcal{B}^{r,p}(X_f) & \xrightarrow{\mathcal{L}^t} & \mathcal{B}^{r,p}(X_f) \end{array}$$

Remark 5.2. It would be preferable if we let $t = 0$ and defined the operator \mathbf{I} as the left inverse of Π . This may be possible but will not be easy.

Let $\beta : S^1 \times \mathbb{R} \rightarrow [0, 1]$ be a smooth function defined by

$$\beta(x, y) = \begin{cases} \chi(\delta_0^{-1}(y - f(x) - 5\delta_0) + 1), & \text{if } f(x) + 5\delta_0 \leq y; \\ 1, & \text{if } 6\delta_0 < y < f(x) + 5\delta_0; \\ 1 - \chi(\delta_0^{-1}(y - 5\delta_0) + 1), & \text{if } y \leq 6\delta_0 \end{cases}$$

where χ is the function defined in (21). This function is taken so that it satisfies

$$\beta(x, y) = \begin{cases} 0, & \text{on the outside of } \tilde{X}_f; \\ 1, & \text{when } 6\delta_0 \leq y \leq f(x) + 5\delta_0 \end{cases}$$

and also

$$\beta(x, f(x) + y) + \beta(x', y) = 1 \quad \text{for any } x, x' \in S^1 \text{ and } 0 \leq y \leq 6\delta_0.$$

We then take C^∞ functions $h_a : \mathbb{R}^2 \rightarrow [0, 1]$ supported on R for $a \in A$ so that³

$$\sum_a h_a \circ \tilde{\kappa}_a^{-1} \equiv \beta \quad \text{on } S^1 \times \mathbb{R}.$$

For each $u \in C^\infty(X_f)$, we set

$$\tilde{u} : \tilde{X}_f \rightarrow \mathbb{C}, \quad \tilde{u}(x, y) = \begin{cases} (\mathcal{L}^{6\delta_0} u)(x, y), & \text{if } y \leq 6\delta_0; \\ u(x, y - 6\delta_0), & \text{if } y \geq 6\delta_0. \end{cases}$$

Since $(\mathcal{L}^{6\delta_0} u)(x, y) = u(x, y - 6\delta_0)$ when $6\delta_0 \leq y \leq f(x)$, this is a smooth function on \tilde{X}_f . We define the operator \mathbf{I} by

$$(53) \quad \mathbf{I}(u) = (u_a)_{a \in A}, \quad u_a = h_a \cdot (\tilde{u} \circ \tilde{\kappa}_a) \quad \text{for } u \in C^\infty(X_f).$$

This operator extends to a bounded operator $\mathbf{I} : \mathcal{B}^{r,p}(X_f) \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R)$, as we will see in the next paragraph, and makes the diagram (52) commutes.

Next we introduce the operator

$$(54) \quad \mathbf{L}^t := \mathbf{I} \circ \mathcal{L}^{t-6\delta_0} \circ \Pi : \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R)$$

for $t \geq 6\delta_0$. By applying Corollary 4.10 to each component, we see that this is a bounded operator. (See Remark 5.6 for more detail.) Since $\mathbf{L}^{6\delta_0} = \mathbf{I} \circ \Pi$ is bounded in particular, so is $\mathbf{I} : \mathcal{B}^{r,p}(X_f) \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R)$ from the definition of $\mathcal{B}^{r,p}(X_f)$. From (52), the diagrams

$$(55) \quad \begin{array}{ccc} \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) & \xrightarrow{\mathbf{L}^t} & \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \\ \Pi \downarrow & & \Pi \downarrow \\ \mathcal{B}^{r,p}(X_f) & \xrightarrow{\mathcal{L}^t} & \mathcal{B}^{r,p}(X_f) \end{array}$$

and

$$(56) \quad \begin{array}{ccc} \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) & \xrightarrow{\mathbf{L}^t} & \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \\ \mathbf{I} \uparrow & & \mathbf{I} \uparrow \\ \mathcal{B}^{r,p}(X_f) & \xrightarrow{\mathcal{L}^t} & \mathcal{B}^{r,p}(X_f) \end{array}$$

commute for $t \geq 6\delta_0$. In particular, the operator $\mathcal{L}^t : \mathcal{B}^{r,p}(X_f) \rightarrow \mathcal{B}^{r,p}(X_f)$ is bounded provided $t \geq 6\delta_0$. It is not difficult to check that the operators

$$\mathcal{L}^t : \mathcal{B}^{r,p}(X_f) \rightarrow \mathcal{B}^{r,p}(X_f) \quad \text{and} \quad \mathbf{L}^t : \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R),$$

have the same essential spectral radius and their peripheral eigenvalues on the outside of it coincide up to multiplicity.

The operator \mathbf{L}^t for $t \geq 6\delta_0$ is expressed as a matrix of operators

$$(57) \quad \mathbf{L}^t(u_a)_{a \in A} = \left(\sum_{a \in A} \mathcal{L}_{a \rightarrow b}^t u_a \right)_{b \in A}.$$

³Here and henceforth, we suppose that $h_a \circ \tilde{\kappa}_a^{-1}$ is a C^∞ function on $S^1 \times \mathbb{R}$ which takes value 0 on the outside of $\tilde{\kappa}_a(Q)$.

Each component $\mathcal{L}_{a \rightarrow b}^t : \mathcal{B}^{r,p}(R) \rightarrow \mathcal{B}^{r,p}(R)$ is written in the form (31), i.e. $\mathcal{L}_{a \rightarrow b}^t u = (\varphi \cdot u) \circ A$ with

$$(58) \quad A = A_{a \rightarrow b}^t : R_{a \rightarrow b}^t \rightarrow \mathbb{R}^2 \quad \text{and} \quad \varphi = \varphi_{a \rightarrow b}^t(x, y) := h_b \circ A_{a \rightarrow b}^t(x, y)$$

where

$$(59) \quad A_{a \rightarrow b}^t(x, y) = \kappa_b^{-1} \circ T_f^t \circ \kappa_a(x, y)$$

and

$$(60) \quad R_{a \rightarrow b}^t = \{z \in R \mid T_f^t \circ \kappa_a(z) \in \kappa_b(R)\}.$$

Remark 5.3. The mapping $A_{a \rightarrow b}^t$ is defined only on a relatively small open subset $R_{a \rightarrow b}^t$ in R , which will be fragmentary in the direction transversal to the flow when t is large. It is locally written in the form (30) with $E \geq 1$ and with g a C^∞ function satisfying $|g'(x)| \leq \gamma_0 \theta_0$. Though the function $\varphi_{a \rightarrow b}^t$ is defined only on $R_{a \rightarrow b}^t$, we may extend it to a C^∞ function on \mathbb{R}^2 with support contained in Q , by letting $\tilde{h} : \mathbb{R}^2 \rightarrow [0, 1]$ be a C^∞ function such that

$$\tilde{h}(z) = \begin{cases} 1, & \text{on } R; \\ 0, & \text{on the outside of } Q, \end{cases}$$

and setting

$$\varphi_{a \rightarrow b}^t(z) = \begin{cases} \tilde{h}(z) \cdot h_b \circ A_{a \rightarrow b}^t(z), & \text{if } z \in Q \text{ and } T_f^t \circ \kappa_a(z) \in \kappa_b(R); \\ 0, & \text{otherwise.} \end{cases}$$

In particular, $\mathcal{L}_{a \rightarrow b}^t$ is smooth on R in the sense that $\mathcal{L}_{a \rightarrow b}^t(C_0^\infty(R)) \subset C_0^\infty(R)$.

5.2. Essential operator norm. We introduce the notion of essential operator norm of a bounded operator. This notion is particularly convenient in our argument about the essential spectral radius. For a bounded operator $L : B \rightarrow B'$ between Banach spaces B and B' , its essential operator norm, denoted by $\|L : B \rightarrow B'\|_{\text{ess}}$, is the infimum of the operator norms of its perturbations by compact operators:

$$\|L : B \rightarrow B'\|_{\text{ess}} := \inf\{\|L - K : B \rightarrow B'\| \mid K : B \rightarrow B' \text{ is compact}\}.$$

Obviously this is bounded by the operator norm $\|L : B \rightarrow B'\|$. Since composition of a compact operator with a bounded operator is again compact, we have

$$\|L' \circ L : B \rightarrow B''\|_{\text{ess}} \leq \|L' : B' \rightarrow B''\|_{\text{ess}} \cdot \|L : B \rightarrow B'\|_{\text{ess}}.$$

The essential spectral radius of $L : B \rightarrow B$ is bounded by its essential norm:

$$\rho_{\text{ess}}(L|_B) \leq \|L^n|_B\|_{\text{ess}}^{1/n} \leq \|L|_B\|_{\text{ess}}.$$

Theorem 3.5 will follow from the claim that, if $\varepsilon > 0$ and if f is sufficiently close to $f_0 \in \mathcal{G}$, there exists some $t_* \geq 6\delta_0$ such that

$$(61) \quad \left\| \mathbf{L}^{t_*} |_{\bigoplus_{a \in A} \mathcal{B}^{r,p,q}(R)} \right\|_{\text{ess}} \leq \exp((\mu(f) + \varepsilon)t_*).$$

Indeed, from Corollary 4.10, we have, for some $C > 0$, that

$$(62) \quad \left\| \mathbf{L}^t |_{\bigoplus_{a \in A} \mathcal{B}^{r,p,q}(R)} \right\| \leq C \quad \text{for } 6\delta_0 \leq t \leq t_* + 6\delta_0.$$

(See Remark 5.6 for more detail.) Since

$$\begin{aligned} \rho_{\text{ess}}(\mathcal{L}^t|_{\mathcal{B}^{r,p}(X_f)}) &\leq \|\mathcal{L}^{nt}|_{\mathcal{B}^{r,p}(X_f)}\|_{\text{ess}}^{1/n} = \|\Pi \circ \mathbf{L}^{t-6\delta_0} \circ \mathbf{I}\|_{\text{ess}}^{1/n} \\ &\leq \|\Pi\|^{1/n} \cdot \|\mathbf{L}^{t*}\|_{\text{ess}}^{\lfloor (nt-6\delta_0)/t_* \rfloor / n} \cdot \|\mathbf{L}^{nt-\lfloor (nt-6\delta_0)/t_* \rfloor \cdot t_*}\|^{1/n} \cdot \|\mathbf{I}\|^{1/n}, \end{aligned}$$

we obtain the conclusion of Theorem 3.5 by letting $n \rightarrow \infty$. In the following subsections, we prove the claim (61).

5.3. Reduction of the claim. Below we show that the claim (61) follows from the corresponding estimates on some localized transfer operators on local charts, to which we can apply Proposition 4.7 and 4.9. We proceed in a few steps. First note that the claim (61) follows if we show that

$$(63) \quad \|\mathcal{L}_{a \rightarrow b}^t : \mathcal{B}^{r,p}(R) \rightarrow \mathcal{B}^{r,p}(R)\|_{\text{ess}} \leq C_0 \exp((\mu(f) + \varepsilon)t)$$

for sufficiently large $t > 0$ and for all $a, b \in A$, with C_0 a constant independent of t . To proceed, we take a finite family of C^∞ functions

$$\{\rho_j^t : \mathbb{R}^2 \rightarrow [0, 1]\}_{j=1}^{J(t)}$$

for each $t > 0$, such that $\sum_{j=1}^{J(t)} \rho_j^t \equiv 1$ on R and that $\text{supp } \rho_j^t \subset Q$. We assume that

- ρ_j^t satisfies (32) with some constants $K_m > 0$ uniform in j and t ,
- the support ρ_j^t is contained in a region of the form $I_j^t \times \mathbb{R}$ where I_j^t is a closed interval on \mathbb{R} , and
- the intersection multiplicity of I_j^t , $1 \leq j \leq J(t)$, is bounded by 2 (say).

Remark 5.4. In the following subsections, we will assume that the length of the interval I_j^t is very small when t is large. It is important that the constants denoted by C_0 below do not depend on the choice of the functions ρ_j^t (though they may depend on the constants K_m).

Let us write the operator $\mathcal{L}_{a \rightarrow b}^t$ as

$$(64) \quad \mathcal{L}_{a \rightarrow b}^t = \sum_{j=1}^{J(t)} \mathcal{M}(\rho_j^t) \circ \mathcal{L}_{a \rightarrow b}^t$$

where $\mathcal{M}(\rho_j^t)$ denotes the multiplication operator by ρ_j^t . By the second claim of Lemma 4.8, we see that the inequality (63) follows if we prove

$$(65) \quad \|\mathcal{M}(\rho_j^t) \circ \mathcal{L}_{a \rightarrow b}^t : \mathcal{B}^{r,p}(R_{a \rightarrow b}^t) \rightarrow \mathcal{B}^{r,p}(R)\|_{\text{ess}} \leq C_0 \exp((\mu(f) + \varepsilon)t)$$

for all $1 \leq j \leq J(t)$, $a, b \in A$ and for sufficiently large $t > 0$, with a constant C_0 independent of t and j .

We decompose the operator $\mathcal{M}(\rho_j^t) \circ \mathcal{L}_{a \rightarrow b}^t$ in (65) further. For $w \in (T_f^t)^{-1}(b)$, there is a unique open neighborhood $U_{b,w}^t$ of the point $w + (0, 6\delta_0)$ in $S^1 \times \mathbb{R}_+$ that is mapped bijectively onto $\kappa_b(R)$ by $T_f^t \circ \pi$. (Recall (50) for the definition of π .) We define

$$R_{a \rightarrow b,w}^t := \kappa_a^{-1}(\pi(U_{b,w}^t)) \cap R \subset R_{a \rightarrow b}^t \subset \mathbb{R}^2.$$

Then $R_{a \rightarrow b}^t$ is the disjoint union of $R_{a \rightarrow b,w}^t$ for $w \in (T_f^t)^{-1}(b)$ though some of $R_{a \rightarrow b,w}^t$ will be empty. Correspondingly we define

$$A_{a \rightarrow b,w}^t = A_{a \rightarrow b}^t|_{R_{a \rightarrow b,w}^t} : R_{a \rightarrow b,w}^t \rightarrow R$$

and, for $1 \leq j \leq J(t)$,

$$\rho_{a \rightarrow b, w, j}^t : R_{a \rightarrow b, w}^t \rightarrow [0, 1], \quad \rho_{a \rightarrow b, w, j}^t(z) = (h_b \cdot \rho_j^t) \circ A_{a \rightarrow b, w}.$$

Then the operator $\mathcal{M}(\rho_j^t) \circ \mathcal{L}_{a \rightarrow b}^t$ is written as the sum

$$\mathcal{M}(\rho_j^t) \circ \mathcal{L}_{a, b}^t = \sum_{w \in (T_f^t)^{-1}(b)} \mathcal{L}_{a \rightarrow b, w, j}^t : C_0^\infty(R) \rightarrow C_0^\infty(R)$$

where $\mathcal{L}_{a \rightarrow b, w, j}^t = 0$ if $R_{a \rightarrow b, w}^t = \emptyset$ and otherwise

$$\mathcal{L}_{a \rightarrow b, w, j}^t u = (\rho_{a \rightarrow b, w, j}^t \cdot u) \circ (A_{a \rightarrow b, w}^t)^{-1}.$$

Remark 5.5. Notice that the functions $\rho_{a \rightarrow b, w, j}^t$ satisfy the condition (32) with some constants $K_m > 0$ uniform for a, b, w, j and t .

By the first claim of Lemma 4.8, the claim (63) follows if we show that

$$(66) \quad \left\| \sum_{w \in (T_f^t)^{-1}(b)} \mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R) \right\|_{\text{ess}} \leq C_0 \exp((\mu(f) + \varepsilon)t)$$

for sufficiently large t and for all $a, b \in A$ and $1 \leq j \leq J(t)$, with a constant C_0 independent of t, a, b and j .

Remark 5.6. Letting the lengths of the intervals I_j^t in the definition of ρ_j^t be small, we may apply Corollary 4.10 to each component $\mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R)$ and see that they are bounded. Consequently the operator \mathbf{L}^t in (57) is bounded. Further, since we may take the bound on the operator norm of \mathbf{L}^t locally uniformly in t , we obtain (62).

5.4. A preliminary argument for the Proof of Theorem 3.5. As we noted in the beginning of this section, in order to illustrate the main point of the argument clearly, we first prove the conclusion of Theorem 3.5 under a stronger assumption. For $n \geq 1$ and $\varepsilon > 0$, we define $\mathcal{G}'(J_\nu, n, \varepsilon; p)$ as the set of $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$ such that, for sufficiently large $t > 0$ and for any $z = (x, y) \in X_f$, the condition (24) holds for any $\xi \in [-\theta_0, \theta_0]$ with $\mathcal{E} = \emptyset$ in the summation. We assume that the roof function f belongs to the set

$$\mathcal{G}' = \bigcap_{\nu=1}^{\nu_0} \bigcap_{m=1}^{\infty} \bigcap_{n \geq 1} \mathcal{G}'(J_\nu, n, 1/m; p) \subset \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0).$$

Remark 5.7. From the discussion preceding to Theorem 3.5, we expect that the subset \mathcal{G}' above is also prevalent in $\mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$. The proof of Theorem 3.5 would be simpler if this was true, as we will see below. But some technical difficulties (related to interference between perturbations) prevent us from this. We therefore resort to a more involved argument presented in the next subsection.

We continue the argument in the last subsection under the assumption as above. We assume that the lengths of the intervals I_j^t in the choice of the functions ρ_j^t are very small. (The precise condition will be given in Remark 5.8.) Let us take and fix a point⁴ $z_0 = z_0(j) \in \text{supp } \rho_j^t \cap R$. For each $w \in (T_f^t)^{-1}(b)$ with $R_{a \rightarrow b, w}^t \neq \emptyset$, let

⁴We will ignore j 's such that $\text{supp } \rho_j^t \cap R = \emptyset$.

$q = q(w) \in Q$ be the unique point satisfying $\kappa_a(q(w)) \in U_{b,w}^t$ and $T_f^t(\kappa_a(q(w))) = \kappa_b(z_0)$. Then let $S(w)$ and $E(w) \geq 1$ be real numbers such that

$$(67) \quad (DA_{a \rightarrow b, w}^t)_{q(w)} = \begin{pmatrix} E(w) & 0 \\ -S(w)E(w) & 1 \end{pmatrix}.$$

We divide the set $(T_f^t)^{-1}(b)$ into disjoint subsets B_ν , $1 \leq \nu \leq \nu_0$, so that $w \in (T_f^t)^{-1}(b)$ is contained in B_ν only if $E(w) \in [e^{a_\nu t}, e^{b_\nu t}]$. Further, letting t be sufficiently large, we may and do suppose that $B_\nu = \emptyset$ if

$$(68) \quad [\chi_{\min}(f), \chi_{\max}(f)] \cap \text{int} J_\nu = \emptyset.$$

Then the operator on the left hand side of (66) is expressed as

$$\sum_{w \in (T_f^t)^{-1}(b)} \mathcal{L}_{a \rightarrow b, w, j}^t = \sum_{\nu=1}^{\nu_0} \Psi_\nu \circ \Phi_\nu$$

where

$$\Phi_\nu : \mathcal{B}^{r,p}(R) \rightarrow \bigoplus_{w \in B_\nu} \mathcal{B}_{S(w), E(w)}^{r,p}(R), \quad \Phi_\nu(u) = (\mathcal{L}_{a \rightarrow b, w, j}^t u)_{w \in B_\nu}$$

and

$$\Psi_\nu : \bigoplus_{w \in B_\nu} \mathcal{B}_{S(w), E(w)}^{r,p}(R) \rightarrow \mathcal{B}^{r,p}(R), \quad \Psi_\nu((u_w)_{w \in B_\nu}) = \sum_{w \in B_\nu} u_w.$$

Here we suppose that $\bigoplus_{w \in B_\nu} \mathcal{B}_{S(w), E(w)}^{r,p}(R)$ is equipped with the norm

$$\|(u_w)\| := \left(\sum_{w \in B_\nu} \|u_w\|_{r,p, S(w), E(w)}^{2p} \right)^{1/2p}$$

Then, from Lemma 4.8 and Proposition 4.7, the essential operator norm of Φ_ν is bounded by

$$(69) \quad C_0 \max_{w \in B_\nu} E(w)^{1/2p} \leq C_0 \exp(b_\nu t / (2p)).$$

Remark 5.8. To get the estimate (69), we apply Proposition 4.7 to each $\mathcal{L}_{a \rightarrow b, w, j}^t$. For this purpose, we have to assume that the lengths of intervals I_j^t in the choice of the functions ρ_j^t are sufficiently small. This is of course possible. The point is that the constant denoted by C_0 in (69) does not depend on the choice of ρ_j^t .

From Proposition 4.9 and (18), the essential operator norm of Ψ_ν is bounded by

$$C_0 \exp((h(f) + \varepsilon)t(p-1)/2p) \cdot \Delta_\nu^{1/2p}$$

where Δ_ν is the quantity defined in Proposition 4.9 in the setting

$$\{(S(i), E(i)) \mid i = 1, \dots, M := \#B_\nu\} = \{(S(w), E(w)) \mid w \in B_\nu\}.$$

Remark 5.9. To deduce the estimate above, we used (18) to bound $\#B_\nu$. Note also that, from the condition (16) in the choice of r , the latter factor $M^{p-1}\Delta \geq M^{p-1}$ on the right hand side of the inequality (44) of Proposition 4.9 exceeds the former factor $M^{2p-1}/(\min_{1 \leq i \leq M} E(i))^{2pr} \lesssim \exp(-2pr \cdot \chi_{\min} t) M^{2p-1}$.

Since we are assuming that $f \in \mathcal{G}'$, we have that

$$\Delta_\nu \leq \exp((\max\{ph(f) - a_\nu, 0\} + p(b_\nu - a_\nu) + \varepsilon)t)$$

for sufficiently large t , uniformly in $a, b \in A$ and $1 \leq j \leq J(t)$. Therefore we conclude that the essential operator norm of $\Psi_\nu \circ \Phi_\nu$ is bounded by

$$\exp\left(\frac{b_\nu + (p-1)(h(f) + \varepsilon) + \max\{ph(f) - a_\nu, 0\} + p(b_\nu - a_\nu) + 2\varepsilon}{2p} \cdot t\right)$$

provided that t is sufficiently large. By the definition of $\mu(f)$ in (26) and arbitrariness of $\varepsilon > 0$, this implies (66). (Note that $\Psi_\nu \circ \Phi_\nu = 0$ if (68) holds.)

5.5. Proof of Theorem 3.5. We explain how we modify the argument in the last subsection in order to get the same conclusion under the weaker assumption of Theorem 3.5. The idea is not difficult: we use the fact that the exceptional set \mathcal{E} is relatively small as we formulate in (71) below.

We resume the argument in Subsection 5.3. Recall that we are considering an arbitrarily small number $\varepsilon > 0$. Let m and m' be large integers that we will specify in the course of the argument. We take $n \geq n_0(1/m)$ so large that

$$(70) \quad \nu_0 p \cdot \lceil 10m\bar{\chi}_{\max} \rceil \leq \exp(\varepsilon n).$$

In the following we assume that f belongs to

$$\bigcap_{\nu=1}^{\nu_0} \mathcal{G}(J_\nu, n, 1/m, 1/m'; p).$$

We take $t_0 > 0$ so that the conditions in the definitions of $\mathcal{G}(J_\nu, n, 1/m, 1/m'; p)$ for $\nu = 1, \dots, \nu_0$ hold for $t \geq t_0$. From the definition of $\mathcal{G}(\cdot)$ in Theorem 3.3, this implies that, for any $t \geq t_0$, any $z = (x, y) \in X_f$ with $x \notin \text{Per}_{1/m'}(\tau, n)$ and $1 \leq \nu \leq \nu_0$, there exists a subset $\mathcal{E} = \mathcal{E}_\nu(z, t; f) \subset \tau^{-n}(x)$ with $\#\mathcal{E} \leq p\lceil 10ma_\nu \rceil$ such that the condition (24) holds for any $\xi \in [-\theta_0, \theta_0]$ with $J = J_\nu$. We put

$$\mathcal{E}(z, t; f) = \bigcup_{\nu=1}^{\nu_0} \mathcal{E}_\nu(z, t; f).$$

From the condition (70) in the choice of n , we have

$$(71) \quad \#\mathcal{E}(z, t; f) \leq \exp(\varepsilon n).$$

We may and do assume further that t_0 is so large that $t_0 > 2n \cdot y_{\max}$ and also

$$(72) \quad \frac{1}{t} \log |\det DT_f^t(w)| \in [\chi_{\min}(f) - \varepsilon, \chi_{\max}(f) + \varepsilon] \quad \text{for any } w \in X_f \text{ and } t \geq t_0.$$

We prove that (66) holds for all $a, b \in A$ and $1 \leq j \leq J(t)$ if $t \geq t_0$ is sufficiently large. (Notice that the constant C_0 in (66) have to be uniform for a, b, t and j .) Suppose $t \geq t_0$ and consider arbitrary $a, b \in A$ and $1 \leq j \leq J(t)$. We take a point $z_0 = z_0(j) \in \text{supp } \rho_j^t \cap R$ and write $\kappa_b(z_0) = (x_0, y_0)$. For each point $x \in \tau^{-k}(x_0)$, we define

$$t(k, x) = f^{(kn)}(x) + y_0$$

so that $T^{t(k, x)}(x, 0) = (x_0, y_0) = \kappa_b(z_0)$ for $x \in \tau^{-kn}(x_0)$. Then, we construct the subsets $H_k \subset \tau^{-kn}(x_0)$ for $k \geq 0$ inductively as follows. For $k = 0$, we set $H_0 = \{x_0\}$. If H_{k-1} for $k \geq 1$ has been defined, let H_k be the set of points $x \in \tau^{-kn}(x_0)$ satisfying

$$(H1) \quad x' := \tau^n(x) \in \tau^{-(k-1)n}(x_0) \text{ belongs to } H_{k-1},$$

$$(H2) \quad t - t(k-1, x') > t_0, \text{ and}$$

(H3) (a) $x' \in \text{Per}_{1/m'}(\tau, n)$, or else (b) $x \in \mathcal{E}((x', 0), t - t(k - 1, x'); f)$.

Remark 5.10. The condition (H2) ensures that the subset $\mathcal{E}((x', 0), t - t(k - 1, x); f)$ in the condition (H3) is well-defined.

We check that the number of points in $H_k \subset \tau^{-kn}(x_0)$ is relatively small compared with $\#\tau^{-kn}(x_0) = \ell^{kn}$. Let us say that $x \in H_{k+\nu}$ is a descendant of ν -th generation of $x' \in H_k$ if $\tau^{\nu n}(x) = x'$. Observe that

- (1) if $x' \notin \text{Per}_{1/m'}(\tau, n)$, the number of its descendants of the first generation is bounded by $\exp(\varepsilon n)$ from (71), and
- (2) if $x' \in \text{Per}_{1/m'}(\tau, n)$, the number of its descendants of the first generation is ℓ^n .

In the case (2) above, the bound on the number of descendants is not effective. But, if the case (2) happens for x' , the same will not happen for most of its descendants for several generations. More precisely, for arbitrarily large $\nu_0 > 0$, we may let m' be so large (depending on n) that the descendant of x' of ν -th generation with $\nu \leq \nu_0$ is not contained in $\text{Per}_{1/m'}(\tau, n)$ but for at most one exception. Therefore, letting m' be large, we may suppose

$$(73) \quad \#H_k \leq C_0 \ell^n \exp(2\varepsilon kn) \quad \text{for } k \geq 0$$

where $C_0 > 0$ is a constant depending only on ε .

Let \mathcal{H} be the set of pairs (k, x) of an integer $k \geq 0$ and a point $x \in H_k$. We say that a pair $(k, x) \in \mathcal{H}$ is *terminal* if $t - t(k, x) \leq t_0$ and write $\mathcal{H}_{\text{term}} \subset \mathcal{H}$ for the set of such pairs. If a pair $(k, x) \in \mathcal{H}$ is terminal, there is no descendant of $x \in H_k$.

Using the definitions prepared above, we divide the set $(T_f^t)^{-1}(b)$ into several (disjoint) subsets. For each $w \in (T_f^t)^{-1}(b)$ with $R_{a \rightarrow b, w}^t \neq \emptyset$, let $q(w) \in R_{a \rightarrow b, w}^t$ be the point such that

$$\tilde{q}(w) := \kappa_a(q(w)) \in U_{b, w}^t \quad \text{and} \quad T_f^t(\tilde{q}(w)) = \kappa_b(z_0).$$

For each $(k, x) \in \mathcal{H}$, let $Q(k, x)$ be the set of points $w \in (T_f^t)^{-1}(b)$ with $R_{a \rightarrow b, w}^t \neq \emptyset$ such that

$$(74) \quad T_f^{s_{kn}(\tilde{z}_0, \tilde{q}(w); t)}(\tilde{q}(w)) = (x, 0) \quad \text{but} \quad T_f^{s_{(k+1)n}(\tilde{z}_0, \tilde{q}(w); t)}(\tilde{q}(w)) \notin H_{k+1} \times \{0\}$$

where $\tilde{z}_0 = (x_0, y_0) = \kappa_b(z_0)$. (Recall (19) for the definition of $s_k(z, w; t)$.) Clearly the set $(T_f^t)^{-1}(b)$ splits into the disjoint subsets $Q(k, x)$ for $(k, x) \in \mathcal{H}$, provided that we ignore $w \in (T_f^t)^{-1}(b)$ with $R_{a \rightarrow b, w}^t = \emptyset$.

Remark 5.11. The former condition in (74) implies that

$$s_{(k-1)n}(\tilde{z}_0, \tilde{q}(w); t) = t - t(k - 1, \tau^n(x)) > t_0 > 2n \cdot y_{\max}$$

and hence that $s_{(k+1)n}(\tilde{z}_0, \tilde{q}(w); t)$ in the latter condition is well-defined.

In the case where a pair $(k, x) \in \mathcal{H}$ is terminal, we have $t - t(k, x) \leq t_0$ by definition and we have $T_f^{t-t(k, x)}(\tilde{q}(w)) = (x, 0)$. In particular, we have

$$(75) \quad \#Q(k, x) \leq \ell^{t_0/y_{\min}} \quad \text{if } (k, x) \in \mathcal{H}_{\text{term}}.$$

In the case where a pair $(k, x) \in \mathcal{H}$ is *not* terminal, we decompose the subset $Q(k, x) \subset (T_f^t)^{-1}(b)$ further. In this case, we have $t - t(k, x) > t_0$. Further, from the definition of H_k 's, we have, for $w \in Q(k, x)$, that

$$T_f^{s_{kn}(\tilde{z}_0, \tilde{q}(w); t)}(\tilde{q}(w)) = T_f^{t-t(k, x)}(\tilde{q}(w)) = (x, 0) \notin \text{Per}_{1/m'}(\tau, n)$$

and

$$T_f^{s_{(k+1)n}(\tilde{z}_0, \tilde{q}(w); t)}(\tilde{q}(w)) = (\tilde{x}, 0) \quad \text{with } \tilde{x} \notin \mathcal{E}((x, 0), t - t(k, x); f).$$

From (72), we can divide $Q(k, x)$ into disjoint subsets $Q_\nu(k, x)$ for $1 \leq \nu \leq \nu_0$ so that $w \in Q(k, x)$ belongs to $Q_\nu(k, x)$ only if

$$(76) \quad \det(DT_f^{t-t(k, x)}(\tilde{q}(w))) \in [e^{a_\nu(t-t(k, x))}, e^{b_\nu(t-t(k, x))}]$$

and also that $Q_\nu(k, x) = \emptyset$ if (68) holds.

We now estimate the essential operator norm of the operator on the left hand side of the claim (66). We decompose the operator into several parts, correspondingly to the decomposition of $(T_f^t)^{-1}(b)$ into $Q_\nu(k, x)$ for $(k, x) \in \mathcal{H}$, and estimates the essential operator norms of those parts. (We ignore $w \in (T_f^t)^{-1}(b)$ with $R_{a \rightarrow b, w}^t = \emptyset$ since $\mathcal{L}_{a \rightarrow b, w, j}^t$ vanishes for such w .)

In general, we have

$$\|\mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R)\|_{\text{ess}} \leq C_0 e^{(\chi_{\max}(f) + \varepsilon)t/2p}$$

by Proposition 4.7 and Proposition 4.9 (in the trivial case of $M = 1$ and $\Delta = 1$.) Hence, by a simple estimate using (73) and (75), we obtain

$$\begin{aligned} & \left\| \sum_{(k, x) \in \mathcal{H}_{\text{term}}} \sum_{w \in Q(k, x)} \mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R) \right\|_{\text{ess}} \\ & \leq C_0 \left(\sum_{k \leq t/(ny_{\min})} \ell^{t_0/y_{\min}} \cdot \ell^n \exp(2\varepsilon kn) \right) \cdot e^{(\chi_{\max}(f) + \varepsilon)t/2p} \end{aligned}$$

where the range of k in the sum on the right hand side is restricted to $k \leq t/(ny_{\min})$ because H_k is empty if $nk \cdot y_{\min} > t$. Since we have

$$(77) \quad \mu(f) \geq \frac{(p-1)h(f) + \chi_{\max}}{2p} > \frac{\chi_{\max}}{2p}$$

from the definition and since we may suppose that $\varepsilon > 0$ is small, we see that the right hand side of the inequality above is bounded by $e^{(\mu(f) + \varepsilon)t}$ when t is sufficient large.

We next consider $(k, x) \in \mathcal{H}$ which is *not* terminal. First of all, observe that, for the case of $(0, x_0) \in H_0$, the argument in the last subsection applies to

$$\sum_{w \in Q(0, x_0)} \mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R)$$

and we can conclude that the essential operator norm of this operator is bounded by $C_0 \exp((\mu + \varepsilon)t)$. (Note that the subset $Q(0, x_0)$ by definition does not contain the problematic elements $w \in (T_f^t)^{-1}(b)$ such that $T_f^{s_n(\tilde{z}_0, \tilde{q}(w); t)}(\tilde{q}(w)) \in \mathcal{E}(\tilde{z}_0, t; f)$.) Below we see that a similar argument applies to the case $k > 0$.

Suppose that $(k, x) \in \mathcal{H}$ is not terminal and $w \in Q(k, x)$. We consider the local chart $\kappa_c : Q \rightarrow X_f$ for $c = (x, 0) \in X_f$ so that the point $(x, 6\delta_0)$ belongs to $\kappa_c(R)$. Let $V \subset Q$ be the neighborhood of $(0, 6\delta_0)$ that is mapped by $T_f^{t(k, x) - 6\delta_0} \circ \kappa_c$ bijectively onto $\kappa_b(R)$. We define

$$E(w) := E(\tilde{q}(w), t - t(k, x); f) := \det(DT_f^{t-t(k, x)}(\tilde{q}(w))) \geq 1$$

and

$$S(w) := -F(\tilde{q}(w), t - t(k, x); f) / E(\tilde{q}(w), t - t(k, x); f)$$

where $E(\cdot)$ and $F(\cdot)$ on the right hand sides are those in (11), so that

$$(78) \quad (DA_{a \rightarrow c}^{t-t(k,x)+6\delta_0})_{q(w)} = \begin{pmatrix} E(w) & 0 \\ -S(w)E(w) & 1 \end{pmatrix}$$

where $A_{a \rightarrow c}^t : R_{a \rightarrow c}^t \rightarrow R$ is defined by (59) and (60) with b replaced by c . Then we may express the operator

$$\sum_{w \in Q(k, x)} \mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R)$$

as

$$\sum_{w \in Q(k, x)} \mathcal{L}_{a \rightarrow b, w, j}^t = \Xi_{k, x} \circ \left(\sum_{1 \leq \nu \leq \nu_0} \Psi_{k, x, \nu} \circ \Phi_{k, x, \nu} \right)$$

where the operators $\Xi_{k, x}$, $\Psi_{k, x, \nu}$ and $\Phi_{k, x, \nu}$ are defined as follows: The operators

$$\Phi_{k, x, \nu} : \mathcal{B}^{r, p}(R) \rightarrow \bigoplus_{w \in Q_\nu(k, x)} \mathcal{B}_{S(w), E(w)}^{r, p}(V)$$

and

$$\Psi_{k, x, \nu} : \bigoplus_{w \in Q_\nu(k, x)} \mathcal{B}_{S(w), E(w)}^{r, p}(V) \rightarrow \mathcal{B}^{r, p}(V)$$

are respectively analogues of the operators Ψ_ν and Φ_ν considered in the last subsection and defined precisely by

$$\Phi_{k, x, \nu}(u) = \left((\rho_{a \rightarrow b, w, j}^t \cdot u) \circ (A_{a \rightarrow c}^{t-t(k,x)+6\delta_0}|_{R_{a \rightarrow b, w}^t})^{-1} \right)_{w \in Q_\nu(k, x)}$$

and

$$\Psi_{k, x, \nu}((u_w)_{w \in Q_\nu(k, x)}) = \sum_{w \in Q_\nu(k, x)} u_w.$$

On the other hand, we define

$$\Xi_{k, x} : \mathcal{B}^{r, p}(V) \rightarrow \mathcal{B}^{r, p}(R), \quad \Xi_{k, x}u = u \circ (A_{c \rightarrow b}^{t(k, x)-6\delta_0}|_V)^{-1}.$$

For the operator $\sum_{1 \leq \nu \leq \nu_0} \Psi_{k, x, \nu} \circ \Phi_{k, x, \nu}$, the situation is parallel to that considered in the last subsection and hence we can get the estimate

$$\left\| \sum_{1 \leq \nu \leq \nu_0} \Psi_{k, x, \nu} \circ \Phi_{k, x, \nu} : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(V) \right\|_{\text{ess}} \leq C_0 \exp((\mu(f) + \varepsilon)(t(k, x) + 6\delta_0))$$

applying Proposition 4.7, Lemma 4.8 and Proposition 4.9. For the operator $\Xi_{k, x}$, we have the estimate

$$\|\Xi_{k, x}\|_{\text{ess}} \leq C_0 \exp((\chi_{\max}(f) + \varepsilon)(t - t(k, x) - 6\delta_0)/2p)$$

from Proposition 4.7 and Proposition 4.9 (in the trivial case of $M = 1$ and $\Delta = 1$). Hence, noting (77), we obtain

$$\left\| \sum_{w \in Q(k, x)} \mathcal{L}_{a \rightarrow b, w, j}^t : \mathcal{B}^{r, p}(R) \rightarrow \mathcal{B}^{r, p}(R) \right\|_{\text{ess}} \leq C_0 \exp((\mu(f) + \varepsilon)t)$$

provided that $\varepsilon > 0$ is sufficiently small. Therefore we conclude (66) by summing these estimates for $(k, x) \in \mathcal{H} \setminus \mathcal{H}_{\text{term}}$ and using (73) and arbitrariness of $\varepsilon > 0$.

We have proved that the conclusion of Theorem 3.5 holds for $f \in \mathcal{G}$. But notice that, to get the conclusion of Theorem 3.5 for some $\eta > 0$, it is actually enough to show the estimate (66) for $\varepsilon = \eta/2$ and some (very) large t according to η and C_0 . Hence the conclusion remains true for small perturbations of f . This completes the proof of Theorem 3.5.

6. PROOF OF THEOREM 2.3

In this section, we justify the heuristic argument in Subsection 2.2 and relate the distribution of the periods of prime periodic orbits of T_f^t and spectral properties of the transfer operators \mathcal{L}^t . In principle, we follow the idea presented in the paper [1] where a similar statement for hyperbolic diffeomorphisms is proved.

6.1. A lift of the operator \mathbf{L}^t . We first introduce a kind of lift (or extension) \mathbb{I}^t of the transfer operator \mathcal{L}^t . Let us recall the definitions of the Banach spaces $\mathcal{B}^{r,p}(\mathbb{R}^2)$ and $\mathcal{B}^{r,p}(X_f)$ in Section 4 and 5 respectively. Using the notation appeared in those definitions, we define the operators

$$\mathbb{I} : \bigoplus_{a \in A} C^\infty(R) \rightarrow \bigoplus_{a \in A} \prod_{m,n} \mathcal{S}(\mathbb{R}^2), \quad \mathbb{I}^* : \bigoplus_{a \in A} \bigoplus_{m,n} \mathcal{S}(\mathbb{R}^2) \rightarrow \bigoplus_{a \in A} C^\infty(\mathbb{R}^2)$$

by

$$\mathbb{I}((u_a)_{a \in A}) = (\hat{\chi}_{m,n}(u_a))_{a \in A, m \in \mathbb{Z}_+, n \in \mathbb{Z}}$$

and

$$\mathbb{I}^*((u_{a,m,n})_{a \in A, m \in \mathbb{Z}_+, n \in \mathbb{Z}}) \mapsto \left(u_a := \sum_{m,n} \hat{\chi}'_{m,n}(u_{a,m,n}) \right)_{a \in A}$$

where $\hat{\chi}'_{m,n}$ is the function introduced in the proof of Lemma 4.5 and the ranges of the variables m and n are $\mathbb{Z}_{\geq 0}$ and \mathbb{Z} respectively. We have $\mathbb{I}^* \circ \mathbb{I} = \text{Id}$ because $\{\chi_{m,n}\}$ is a partition of unity on \mathbb{R}^2 and $\hat{\chi}'_{m,n} \cdot \chi_{m,n} = \chi_{m,n}$.

Let $\mathbb{B}^{r,p}$ be the Banach space obtained as the completion of the space $\bigoplus_{m,n} \mathcal{S}(\mathbb{R}^2)$ with respect to the norm

$$\|(u_{m,n})\|_{r,p}^{(\varepsilon_0)} = \left(\sum_{m,n} 2^{2rpm} \cdot \varepsilon(m) \cdot \|u_{m,n}\|_{L^{2p}}^{2p} \right)^{1/2p}$$

where

$$\varepsilon(m) = \begin{cases} \varepsilon_0, & \text{if } m = 0; \\ 1, & \text{otherwise} \end{cases}$$

with $\varepsilon_0 > 0$ a small constant that we will specify later. From the definition of the norm on $\mathcal{B}^{r,p}(\mathbb{R}^2)$, we see that the operators \mathbb{I} and \mathbb{I}^* extend to bounded operators

$$\mathbb{I} : \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \rightarrow \bigoplus_{a \in A} \mathbb{B}^{r,p}, \quad \mathbb{I}^* : \bigoplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \bigoplus_{a \in A} \mathcal{B}^{r,p}(R).$$

Remark 6.1. In the definition of the norm $\|\cdot\|_{r,p}^{(\varepsilon_0)}$ above, we put the factor $\varepsilon(m)$ by a technical reason. Of course, the Banach space $\mathbb{B}^{r,p}$ (as a set) does not depend on the choice of the constant $\varepsilon_0 > 0$. If $\varepsilon_0 = 1$, the operator \mathbb{I} is an isometric injection by definition.

For $t \geq 6\delta_0$, we define

$$\mathbb{L}^t : \bigoplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \bigoplus_{a \in A} \mathbb{B}^{r,p}, \quad \mathbb{L}^t = \mathbb{I} \circ \mathbf{L}^t \circ \mathbb{I}^*.$$

Also, for a C^∞ function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ compactly supported on $[6\delta_0, +\infty)$, we define

$$\mathbb{L}^\varphi = \int \varphi(t) \cdot \mathbb{L}^t dt : \bigoplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \bigoplus_{a \in A} \mathbb{B}^{r,p}.$$

Note that the following diagram commutes:

$$\begin{array}{ccc} \bigoplus_{a \in A} \mathbb{B}^{r,p} & \xrightarrow{\mathbb{L}^t} & \bigoplus_{a \in A} \mathbb{B}^{r,p} \\ \mathbb{I} \uparrow & & \mathbb{I} \uparrow \\ \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) & \xrightarrow{\mathbf{L}^t} & \bigoplus_{a \in A} \mathcal{B}^{r,p}(R) \\ \Pi \downarrow & & \Pi \downarrow \\ \mathcal{B}^{r,p}(X_f) & \xrightarrow{\mathcal{L}^t} & \mathcal{B}^{r,p}(X_f) \end{array}$$

The essential spectral radii of the three operators \mathcal{L}^t , \mathbf{L}^t and \mathbb{L}^t in the commutative diagram above are same and their peripheral eigenvalues outside of it coincide. Indeed, we have checked this relation between \mathcal{L}^t and \mathbf{L}^t in Subsection 5.1 and, similarly, can check this relation between \mathbf{L}^t and \mathbb{L}^t .

We define the flat trace of operators on $\bigoplus_{a \in A} \mathbb{B}^{r,p}$. Let $\mathbb{B}_{a,m,n}^{r,p}$ be the Banach space $L^{2p}(\mathbb{R}^2)$ equipped with the norm $\|u\|_m^{(\varepsilon_0)} = 2^{rm} \cdot \varepsilon(m) \cdot \|u\|_{L^{2p}}$. Then a bounded operator $\mathbb{M} : \bigoplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \bigoplus_{a \in A} \mathbb{B}^{r,p}$ may be regarded as a matrix of operators whose components are

$$\mathbb{M}_{(a,m,n) \rightarrow (a',m',n')} : \mathbb{B}_{a,m,n}^{r,p} \rightarrow \mathbb{B}_{a',m',n'}^{r,p}, \quad u \mapsto \pi_{a',m',n'} \circ \mathbb{M} \circ \pi_{a,m,n} u$$

where $\pi_{a,m,n}$ is the projection to the (a, m, n) -component. If all the diagonal components $\mathbb{M}_{(a,m,n) \rightarrow (a,m,n)}$ are trace class operators and if

$$\sum_{a,m,n} |\mathrm{Tr} \mathbb{M}_{(a,m,n) \rightarrow (a,m,n)}| < \infty,$$

we define the flat trace of \mathbb{M} by

$$\mathrm{Tr}^b \mathbb{M} = \sum_{a,m,n} \mathrm{Tr} \mathbb{M}_{(a,m,n) \rightarrow (a,m,n)}.$$

Lemma 6.2. *If \mathbb{M} is a trace class operator, its flat trace is well defined and we have $\mathrm{Tr}^b \mathbb{M} = \mathrm{Tr} \mathbb{M}$.*

Proof. Suppose that \mathbb{M} is a rank one operator of the form $\mathbb{M}u = y(u) \cdot x$ with

$$x = (x_{a,m,n}) \in \bigoplus_{a \in A} \mathbb{B}^{r,p}, \quad y = (y_{a,m,n}) \in \bigoplus_{a \in A} (\mathbb{B}^{r,p})^*.$$

Note that, letting $q > 1$ be such that $q^{-1} + (2p)^{-1} = 1$, we have

$$\|y\|_{(\oplus_{a \in A} \mathbb{B}^{r,p})^*} = \left(\sum_{a,m,n} \|y_{a,m,n}\|_{(\mathbb{B}_{a,m,n}^{r,p})^*}^q \right)^{1/q}$$

where

$$\|v\|_{(\mathbb{B}_{a,m,n}^{r,p})^*} = 2^{-rm} \cdot \varepsilon(m)^{-1} \cdot \|v\|_{L^q}.$$

Hence we have, by Hölder inequality, that

$$\begin{aligned} \sum_{a,m,n} |\mathrm{Tr} \mathbb{M}_{(a,m,n) \rightarrow (a,m,n)}| &\leq \sum_{a,m,n} \|x_{a,m,n}\|_{\mathbb{B}_{a,m,n}^{r,p}} \cdot \|y_{a,m,n}\|_{(\mathbb{B}_{a,m,n}^{r,p})^*} \\ &\leq \|x\|_{\oplus_{a \in A} \mathbb{B}^{r,p}} \cdot \|y\|_{\oplus_{a \in A} (\mathbb{B}^{r,p})^*} \end{aligned}$$

and also that

$$\sum_{a,m,n} \mathrm{Tr} \mathbb{M}_{(a,m,n) \rightarrow (a,m,n)} = \sum_{a,m,n} y_{a,m,n}(x_{a,m,n}) = y(x).$$

This and the definition of the trace norm give the conclusion. \square

The flat trace may be defined for operators that are not of the trace class. We introduce such a class of operators. We say that a bounded linear operator $\mathbb{M} : \oplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \oplus_{a \in A} \mathbb{B}^{r,p}$ is *triangular* if its components $\mathbb{M}_{(a,m,n) \rightarrow (a',m',n')}$ is zero whenever

$$2^m \langle n \rangle^2 \leq 2^{m'} \langle n' \rangle^2.$$

Obviously, if \mathbb{M} is triangular, all of its diagonal components vanish and hence its flat trace is defined to be zero. Note that the sum and composition of two triangular operators are again triangular.

In the next subsection, we will see that the operator \mathbb{L}^φ for a C^∞ function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ compactly supported on $[t_0, +\infty)$ with sufficiently large t_0 is decomposed into a trace class operator and a triangular operator. Hence the flat trace is well defined for \mathbb{L}^φ . Note that the components of \mathbb{L}^φ are written as

$$\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi : \mathbb{B}_{a,m,n}^{r,p} \rightarrow \mathbb{B}_{a',m',n'}^{r,p}, \quad u \mapsto \hat{\chi}_{m',n'} \circ \left(\int \varphi(t) \cdot \mathcal{L}_{a \rightarrow a'}^t dt \right) \circ \hat{\chi}_{m,n}' u.$$

This is an integral operator with kernel $K_{(a,m,n) \rightarrow (a',m',n')}^\varphi(z', z) \in \mathcal{S}(\mathbb{R}^4)$ and hence is a trace class operator. If $(a, m, n) = (a', m', n')$, its trace is calculated as the integration of the kernel on the diagonal:

$$\mathrm{Tr} \mathbb{L}_{(a,m,n) \rightarrow (a,m,n)}^\varphi = \int_{\mathbb{R}^2} K_{(a,m,n) \rightarrow (a,m,n)}^\varphi(z, z) dz.$$

Since $\sum_{m,n: 2^m \langle n^2 \rangle \leq k} \hat{\chi}_{m,n}$ converges to the Dirac function δ_0 as $k \rightarrow \infty$, we find

$$(79) \quad \mathrm{Tr} {}^b \mathbb{L}^\varphi = \sum_{a,m,n} \mathrm{Tr} \mathbb{L}_{(a,m,n) \rightarrow (a,m,n)}^\varphi = \sum_{\gamma \in \Gamma} \sum_{n=1}^{\infty} \frac{|\gamma| \cdot \varphi(n|\gamma|)}{1 - E_\gamma^{-n}}$$

by straightforward computation. (Though this computation is not very simple, we ask the readers to check it.) Note that the right hand side is what we expect for $\int \varphi(t) \mathrm{Tr} {}^b \mathcal{L}^t dt$ from (5).

6.2. A decomposition of the lifted operator. We decompose the operator \mathbb{L}^t for $t \geq t_0$ (resp. \mathbb{L}^φ for $\varphi \in C_0^\infty([t_0, +\infty))$) into two parts as

$$\mathbb{L}^t = \mathbb{L}_{\text{trace}}^t + \mathbb{L}_{\text{trace-free}}^t \quad (\text{resp. } \mathbb{L}^\varphi = \mathbb{L}_{\text{trace}}^\varphi + \mathbb{L}_{\text{trace-free}}^\varphi)$$

where $\mathbb{L}_{\text{trace-free}}^t$ (resp. $\mathbb{L}_{\text{trace-free}}^\varphi$) consists of its components $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t$ (reps. $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi$) that satisfies the condition

$$(80) \quad 2^{m'} \langle n' \rangle^2 < 2^{m+4} \langle n \rangle^2 \exp(-(\chi_{\min} + \varepsilon)t)$$

(resp. the same condition with $t = \min \text{supp } \varphi \geq t_0$) and the operator $\mathbb{L}_{\text{trace-free}}^t$ (resp. $\mathbb{L}_{\text{trace-free}}^\varphi$) consists of the remaining components. Clearly the trace-free part $\mathbb{L}_{\text{trace-free}}^t$ (resp. $\mathbb{L}_{\text{trace-free}}^\varphi$) is triangular provided that t_0 is sufficiently large.

Below we present two lemmas, whose proofs are deferred to the next subsection. The constants $r > 0$ and $t_0 > 0$ are assumed to be sufficiently large. The first lemma tells that the trace-free part is strongly contracting. This is a consequence of the choice of the weights in the definition of the Banach space $\mathbb{B}^{r,p}$ and may be rather obvious.

Lemma 6.3. *For $t_0 \leq t \leq 2t_0$, we have*

$$\left\| \mathbb{L}_{\text{trace-free}}^t : \bigoplus_{a \in A} \mathbb{B}^{r,p} \rightarrow \bigoplus_{a \in A} \mathbb{B}^{r,p} \right\| \leq \exp(\rho t)$$

where $\rho = \rho_p(f)$. (See (8) for the definition of $\rho_p(f)$.)

Remark 6.4. As we will see in the proof, we can actually prove the statement above for arbitrarily small ρ by letting r and t_0 larger.

The next lemma proves that the trace class part $\mathbb{L}_{\text{trace}}^\varphi$ is a trace class operator and also gives an estimates on its trace norm.

Lemma 6.5. *For a bounded subset \mathcal{X} in $C_0^\infty([-1, 1])$, there exists a constant $C_* = C_*(\mathcal{X})$ such that, if φ is supported on $[t_0, 2t_0]$ and if there exists an affine map $A(t) = \alpha t + \beta$ with $\alpha \in (0, 1)$ such that the function $\varphi \circ A(t) = \varphi(\alpha t + \beta)$ belongs to \mathcal{X} , then*

$$\|\mathbb{L}_{\text{trace}}^\varphi\|_{\text{Tr}} \leq C_* \alpha \quad \text{and} \quad \|\mathbb{L}_{\text{trace}}^t \circ \mathbb{L}^\varphi\|_{\text{Tr}} \leq C_* \alpha \quad \text{for } t_0 \leq t \leq 2t_0.$$

Remark 6.6. The operator $\mathbb{L}_{\text{trace}}^t$ itself will not be a trace class operator. In the proof of the lemma above, we use the fact that the integration with respect to the variable t (with multiplication by $\varphi(t)$) in the definition of \mathbb{L}^φ reduces the part of functions that have high frequency in the flow direction (that is, the (a, m, n) -components with $|n|$ large).

We proceed with the proof of Theorem 2.3. Below we consider the situation assumed in Theorem 2.3. We write \mathbb{I} for the spectral projector of \mathbb{L}^t for the set of eigenvalues on the outside of the disk $|z| \leq e^{(\rho+\varepsilon)t}$ where $\rho = \rho_p(f)$. Note that this spectral projector \mathbb{I} is of finite rank and does not depend on t provided $t \geq t_0$. By letting t_0 be larger if necessary, we assume

$$\|\mathbb{I} \circ \mathbb{L}^t : \mathcal{B}^{r,p}(X_f) \rightarrow \mathcal{B}^{r,p}(X_f)\| \leq \exp((\rho + \varepsilon)t) \quad \text{for } t \geq t_0.$$

The next proposition is the key to the proof of Theorem 2.3.

Proposition 6.7. *Suppose that \mathcal{X} is a bounded subset in $C_0^\infty([-1, 1])$ and that φ is a C^∞ function supported on $[t_0, 2t_0]$ such that $\varphi \circ A(t) = \varphi(\alpha t + \beta)$ belongs to \mathcal{X} for some affine map $A(t) = \alpha t + \beta$ with $\alpha \in (0, 1)$. Then we have*

$$(81) \quad |\mathrm{Tr}^b((1 - \mathbb{P}) \circ \mathbb{L}^\varphi \circ \mathbb{L}^T)| \leq C'_* \alpha^{-1} \cdot \exp((\rho + \varepsilon)T) \quad \text{for any } T \geq t_0$$

where the constant $C'_* = C'_*(\mathcal{X}, t_0)$ depends on the bounded subset $\mathcal{X} \subset C^\infty([-1, 1])$, t_0 and ε , but not on α and β .

Proof. Let us write $T \geq t_0$ as a sum $T = t_1 + t_2 + \cdots + t_m$ with $t_0 \leq t_i \leq 2t_0$. Since the operators \mathbb{P} , \mathbb{L}^t and \mathbb{L}^φ commute and since $1 - \mathbb{P}$ is a projection operator, we may write

$$\begin{aligned} (1 - \mathbb{P}) \circ \mathbb{L}^\varphi \circ \mathbb{L}^T &= (1 - \mathbb{P}) \circ \mathbb{L}^{t_m} \circ \mathbb{L}^\varphi \circ (1 - \mathbb{P}) \circ \mathbb{L}^{t_{m-1}} \circ \mathbb{L}^{t_{m-2}} \circ \cdots \circ \mathbb{L}^{t_2} \circ \mathbb{L}^{t_1} \\ &= (\mathbb{L}_{\mathrm{trace}}^{t_m} - \mathbb{P} \circ \mathbb{L}^{t_m}) \circ \mathbb{L}^\varphi \circ (1 - \mathbb{P}) \circ \mathbb{L}^{t_{m-1}} \circ \cdots \circ \mathbb{L}^{t_2} \circ \mathbb{L}^{t_1} \\ &\quad + \mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ [\mathbb{L}^\varphi \circ (1 - \mathbb{P}) \circ \mathbb{L}^{t_{m-1}} \circ \cdots \circ \mathbb{L}^{t_2} \circ \mathbb{L}^{t_1}]. \end{aligned}$$

Applying the same deformation to the operator in the last bracket $[\cdot]$ and continuing this procedure, we express the operator $(1 - \mathbb{P}) \circ \mathbb{L}^\varphi \circ \mathbb{L}^T$ as the sum of

$$\begin{aligned} &\mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_1} \circ \mathbb{L}_{\mathrm{trace-free}}^\varphi, \quad \mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_1} \circ \mathbb{L}_{\mathrm{trace}}^\varphi, \\ &\mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_2} \circ (\mathbb{L}_{\mathrm{trace}}^{t_1} - \mathbb{P} \circ \mathbb{L}^{t_1}) \circ \mathbb{L}^\varphi \end{aligned}$$

and

$$\mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_{k+1}} \circ (\mathbb{L}_{\mathrm{trace}}^{t_k} - \mathbb{P} \circ \mathbb{L}^{t_k}) \circ \mathbb{L}^\varphi \circ (1 - \mathbb{P}) \circ \mathbb{L}^{t_{k-1}} \circ \cdots \circ \mathbb{L}^{t_1}$$

for $k = 2, \dots, m$.

Remark 6.8. At the last stage of the development above, we find the term

$$\begin{aligned} &\mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_2} \circ \mathbb{L}^\varphi \circ (1 - \mathbb{P}) \circ \mathbb{L}^{t_1} \\ &= \mathbb{L}_{\mathrm{trace-free}}^{t_m} \circ \cdots \circ \mathbb{L}_{\mathrm{trace-free}}^{t_2} \circ \mathbb{L}^{t_1} \circ (1 - \mathbb{P}) \circ \mathbb{L}^\varphi. \end{aligned}$$

This term is decomposed into the first three terms above.

Notice that the first operator above is triangular and hence its flat trace vanishes. From Lemma 6.3 and Lemma 6.5, the second operator is a trace class operator and its trace norm is bounded by

$$\exp(\rho(t_1 + \cdots + t_m)) \cdot C_* \alpha^{-1}.$$

Similarly the trace norm of the other operators are bounded by

$$\exp\left(\rho \sum_{i=k+1}^m t_i\right) \cdot (C_* \alpha^{-1} + C) \cdot C \exp\left((\rho + \varepsilon) \sum_{i=1}^{k-1} t_i\right)$$

for $k = 1$ and $k = 2, \dots, m$ respectively, where C is a constant depending only on the choice of t_0 and the rank of \mathbb{P} . Hence the claim (81) follows from the estimates above. \square

Since the spectral projector \mathbb{P} is of finite rank, so is $\mathbb{P} \circ \mathbb{L}^t$ and therefore we have

$$(82) \quad \mathrm{Tr}^b(\mathbb{P} \circ \mathbb{L}^\varphi \circ \mathbb{L}^t) = \mathrm{Tr}(\mathbb{P} \circ \mathbb{L}^\varphi \circ \mathbb{L}^t) = \sum_{i=1}^{k'} \int \varphi(t) e^{\mu_i t} dt.$$

Thus Proposition 6.7 is convincing. Below we give an argument to finish the proof. We first justify the relation (6). Let us put

$$\tilde{\pi}(T) = \sum_{\gamma \in \Gamma} \sum_{n=1}^{\lfloor T/|\gamma| \rfloor} \frac{1}{n} \cdot \frac{1}{1 - E_{\gamma}^{-n}}.$$

Note that, if we defined $\mathrm{Tr}^b \mathcal{L}^t$ by the formula (5), we might write it formally as

$$\tilde{\pi}(T) = \int_{+0}^T \frac{1}{t} \cdot \mathrm{Tr}^b \mathcal{L}^t dt.$$

The difference between $\pi(T)$ and $\tilde{\pi}(T)$ is bounded by

$$\sum_{\gamma \in \Gamma} \sum_{n=2}^{\lfloor T/|\gamma| \rfloor} \frac{n^{-1}}{1 - E_{\gamma}^{-n}} + \sum_{\gamma \in \Gamma: |\gamma| \leq T} \frac{E_{\gamma}^{-1}}{1 - E_{\gamma}^{-1}} \leq \sum_{n=2}^{\infty} \sum_{\gamma \in \Gamma: |\gamma| \leq T/n} 1 + \sum_{\gamma \in \Gamma: |\gamma| \leq T} 2 \cdot E_{\gamma}^{-1}.$$

As we noted in Subsection 2.2 this difference is negligible. By the general argument on the topological pressure of flows (see [7, Theorem C] for instance), we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \sum_{\gamma \in \Gamma: |\gamma| \leq T} 1 = P_{\mathrm{top}}(T_f^t, 0) = h(f)$$

and

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log \sum_{\gamma \in \Gamma: |\gamma| \leq T} E_{\gamma}^{-1} = P_{\mathrm{top}}(T_f^t, -\log \det DT_f^t) = 0.$$

Hence we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \log |\tilde{\pi}(T) - \pi(T)| \leq \frac{h(f)}{2}.$$

That is, the difference between $\pi(T)$ and $\tilde{\pi}(T)$ converges to 0 much faster than the error term in the conclusion of Theorem 2.3. Therefore it is enough to prove the statement with $\pi(T)$ replaced by $\tilde{\pi}(T)$.

In order to the last step of the proof, we introduce a few definitions. Let

$$\mu = (h(f) - \rho)/2.$$

For large $T \gg t_0$, we take C^∞ functions

$$\varphi_i^T : \mathbb{R} \rightarrow [0, 1] \quad \text{for } \lceil t_0 \rceil + 1 \leq i \leq \lfloor T \rfloor$$

and

$$\psi_i^T : \mathbb{R} \rightarrow [0, 1] \quad \text{for } 0 \leq i \leq k(T) := \lceil \mu T / \log 2 \rceil$$

so that

- (i) The supports of φ_i^T and ψ_i^T are contained respectively in the intervals

$$I_i = [i - 1, i + 1] \quad \text{and} \quad J_i = [T - 2^{-i}, T + 2^{-k(T)}].$$

- (ii) Let $A_i : [0, 1] \rightarrow I_i$ and $A'_i : [0, 1] \rightarrow J_i$ be the (unique) orientation preserving affine bijections. Then the set of functions

$$\{\varphi_i^T \circ A_i\}_{1 \leq i \leq \lfloor T \rfloor} \quad \text{and} \quad \{\psi_i^T \circ A'_i\}_{1 \leq i \leq k(T)}$$

are contained in a bounded subset $\mathcal{X} \subset C_0^\infty([0, 1])$ that is independent of T .

(iii) If we put $\Psi_k(t) := \sum_{i=\lceil t_0 \rceil+1}^{\lfloor T \rfloor} \varphi_i^T(t) + \sum_{i=0}^k \psi_i^T(t)$ for $1 \leq k \leq k(T)$, we always have that $0 \leq \Psi_k(t) \leq 1/t$ and that

$$\Psi_{k(T)}(t) = \begin{cases} 1/t, & \text{for } t \in [t_0 + 2, T]; \\ 0, & \text{for } t \leq t_0 \text{ and } t \geq T + 2^{-k(T)} \end{cases}$$

and, for $0 \leq k \leq k(T) - 1$,

$$\Psi_k(t) = \begin{cases} 1/t, & \text{for } t \in [t_0 + 2, T - 2^{-k}]; \\ 0, & \text{for } t \leq t_0 \text{ and } t \geq T. \end{cases}$$

We complete the proof of Theorem 2.3 (assuming Lemma 6.3 and 6.5). From the condition (iii) above and (79), we have that

$$\mathrm{Tr}^b \left(\int \Psi_{k(T)-1}(t) \mathbb{L}^t dt \right) \leq \tilde{\pi}(T) \leq \mathrm{Tr}^b \left(\int \Psi_{k(T)}(t) \mathbb{L}^t dt \right) + \tilde{\pi}(t_0 + 2)$$

Hence the difference $\left| \tilde{\pi}(T) - \mathrm{Tr}^b \left(\int_1^T (1/t)(\mathbb{P} \circ \mathbb{L}^t) dt \right) \right|$ is bounded by

$$\begin{aligned} & \int \psi_{k(T)}(t) |\mathrm{Tr}^b(\mathbb{P} \circ \mathbb{L}^t)| dt \\ & + \sum_{i=\lceil t_0 \rceil+1}^{\lfloor T \rfloor} \left| \mathrm{Tr}^b((1 - \mathbb{P}) \circ \mathbb{L}^{\varphi_i^T}) \right| + \sum_{i=0}^{k(T)} \left| \mathrm{Tr}^b((1 - \mathbb{P}) \circ \mathbb{L}^{\psi_i^T}) \right| \end{aligned}$$

plus a constant independent of T . By the estimate (81), we see that the second and third terms are bounded by

$$\sum_{k=1}^{\lfloor T \rfloor} C'_* \exp((\rho + \varepsilon)k) \quad \text{and} \quad C'_* \sum_{k=0}^{k(T)} \exp((\rho + \varepsilon)T + k \log 2)$$

respectively. Hence their sum is bounded by $C \exp((\rho + \mu + \varepsilon)T)$. The first term is bounded by $C \exp((h(f) - \mu + \varepsilon)T)$ because

$$|\mathrm{Tr}^b(\mathbb{P} \circ \mathbb{L}^t)| \leq \mathrm{rank} \mathbb{P} \cdot \exp((h(f) + \varepsilon)t)$$

for sufficiently large t . Therefore, from the choice of μ , we obtain

$$\left| \tilde{\pi}(T) - \int_{t_0}^T \mathrm{Tr}^b(\mathbb{P} \circ \mathbb{L}^t) dt \right| \leq C \exp((h(f) + \rho + \varepsilon)T/2).$$

Clearly the conclusion of Theorem 2.3 follows from this estimate and (82).

Remark 6.9. In the last part of the argument above, we find the reason for the choice of $\mu = (h(f) - \rho)/2$. This also explains why we had the average $\bar{\rho}$ in the statement of Theorem 2.3.

6.3. Proof of Lemma 6.3 and 6.5. Lemma 6.3 and 6.5 follows from elementary estimates on the components of the operators \mathbb{L}^t and \mathbb{L}^φ . If we let $t_0 > 0$ be sufficiently large, we have the following two lemmas.

Lemma 6.10. *For any $\nu > 0$, there exists a constant $C_\nu > 0$ such that*

$$\|\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t : L^{2p}(R) \rightarrow L^{2p}(R)\| \leq C_\nu \exp((\chi_{\max} + \varepsilon)t/p) \cdot \Delta_1(n, n')^{-\nu}$$

for any $a, a' \in A$, for any integers $n, n', m \geq 0, m' \geq 0$ and for any $t \geq t_0$, where $\Delta_1(n, n')$ is that defined in (35). Further, if

$$m' > 0, \quad 2^{m'} \langle n' \rangle^2 > 2^{m+4} \langle n \rangle^2 \cdot \exp(-(\chi_{\min} - \varepsilon)t) \quad \text{and} \quad t_0 \leq t \leq 2t_0,$$

we have

$$\|\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t : L^{2p}(R) \rightarrow L^{2p}(R)\| \leq C_\nu \cdot \max\{2^m \langle n \rangle^2, 2^{m'} \langle n' \rangle^2\}^{-\nu}.$$

Proof. The claim is proved by inspecting the kernel of $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t$ and using integration by parts. We omit the detail of the proof because the argument is parallel to that in the latter part of the proof of Proposition 4.7. \square

Lemma 6.11. *Let $\mathcal{X} \subset C^\infty([-1, 1])$ be a bounded subset. For any $\nu > 0$, there exists a constant $C_\nu(\mathcal{X})$ such that, if φ is supported on $[t_0, 2t_0]$ and if there exists an affine map $A(t) = \alpha t + \beta$ with $\alpha > 0$ such that the function $\varphi \circ A(t) = \varphi(\alpha t + \beta)$ belongs to \mathcal{X} , then we have*

$$(83) \quad \|\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi : L^{2p}(R) \rightarrow L^{2p}(R)\| \leq C_\nu(\mathcal{X}) \cdot \alpha \cdot \langle \alpha |n|^2 \rangle^{-\nu} \cdot \Delta_1(n, n')^{-\nu}.$$

Proof. We proof is again parallel to that of Proposition 4.7. We write the integral kernel of the operator

$$\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi = \int \varphi(t) \cdot \mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t dt$$

explicitly and apply integration by parts. This time, we apply integration by parts also to the integration with respect to the variable t . (Note that the mapping $A_{a \rightarrow a'}^t$ on local charts satisfies $A_{a \rightarrow a'}^{t+\tau}(x, y) = A_{a \rightarrow a'}^t(x, y + \tau)$ when $|\tau|$ is small.) Then we obtain the factor $\alpha^{-1} \cdot \langle \alpha^{-1} |n|^2 \rangle^{-\nu}$ in addition. \square

From the first claim of Lemma 6.10 and the definition of $\mathbb{L}_{\text{trace-free}}^t$, we obtain Lemma 6.3 provided that we let the constant $\varepsilon_0 > 0$ in the definition of $\|(u_{m,n})\|_{r,p}^{(\varepsilon_0)}$ be sufficiently small and let t_0 be sufficiently large. To prove Lemma 6.5, we note

$$\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t = \hat{\chi}'_{m',n'} \circ \mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t.$$

Recalling Lemma 4.5, we see that this implies the estimate

$$\begin{aligned} \|\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t : L^{2p}(R) \rightarrow L^{2p}(R)\|_{\text{Tr}} \\ \leq C_0 2^{m'} \langle n' \rangle^3 \cdot \|\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t : L^{2p}(R) \rightarrow L^{2p}(R)\| \end{aligned}$$

and the same estimate with $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t$ replaced by $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi$. Hence the estimates on the operator norms of $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^t$ and $\mathbb{L}_{(a,m,n) \rightarrow (a',m',n')}^\varphi$ in Lemma 6.10 and Lemma 6.11 give the corresponding estimates on the trace norm. Finally we evaluate the sum of the trace norms of the components of $\mathbb{L}_{\text{trace}}^\varphi$ and $\mathbb{L}_{\text{trace}}^t \circ \mathbb{L}^\varphi$, by using the estimates thus obtained, and conclude Lemma 6.5. (Though this final step is not completely simple, we omit the detail because the estimates involved are straightforward and crude.)

7. PROOF OF THEOREM 3.3

The proof of Theorem 3.3 presented below is basically in the same line as the corresponding argument in the author's previous paper [17]. But we need to modify the argument in some places.

7.1. Families of roof functions. We consider the family of C^∞ functions

$$(84) \quad f_{\mathbf{s}}(x) = f(x) + \sum_{k=1}^K s_k \cdot g_k(x) \quad \text{with parameter } \mathbf{s} = (s_1, s_2, \dots, s_K)$$

for $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0) \subset C_+^\infty(S^1)$ and C^∞ functions

$$(85) \quad g_k : S^1 \rightarrow \mathbb{R}, \quad 1 \leq k \leq K.$$

The range of parameter will be restricted to

$$R(\sigma) = \{\mathbf{s} = (s_1, s_2, \dots, s_K) \mid |s_k| \leq \sigma \text{ for } 1 \leq k \leq K\}$$

for some small $\sigma > 0$. The choice of the functions $g_k \in C_+^\infty(S^1)$ and the constant $\sigma > 0$ will be given in the course of the argument below.

We consider an interval $J = [a, b]$, as in the statement of Theorem 3.3. We suppose $0 < \varepsilon < \min\{a, 1\}$ and set

$$(86) \quad q = q(\varepsilon) := \left\lceil \frac{10a}{\varepsilon} \right\rceil.$$

Below we consider an integer $n \geq 1$ and show that the conclusion of Theorem 3.3 holds when n is sufficiently large according to ε .

Let $x \in S^1$ and $m \geq 1$. For each p -tuple of points in $\tau^{-mn}(x)$,

$$\mathbf{x} = (\mathbf{x}(i))_{i=1}^p \in (\tau^{-mn}(x))^p,$$

we set

$$(87) \quad \hat{S}(\mathbf{x}, k; f_{\mathbf{s}}) = \ell^{-k} \sum_{i=1}^p \frac{d}{dx} f_{\mathbf{s}}^{(k)}(\mathbf{x}(i)).$$

For an array $X = (\mathbf{x}_1, \dots, \mathbf{x}_q)$ of q elements in $(\tau^{-mn}(x))^p$, we consider the map

$$\Phi_{x,X} : \mathbb{R}^K \rightarrow \mathbb{R}^q, \quad \Phi_{x,X}(\mathbf{s}) = \left(\hat{S}(\mathbf{x}_j, mn; f_{\mathbf{s}}) \right)_{j=1}^q.$$

This is an affine map and its linear part does not depend on $f = f_0$.

Definition 7.1. We say that an (ordered) array of q elements in $(\tau^{-n}(x))^p$,

$$(88) \quad X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_q)$$

is *independent* if there is a component $\mathbf{x}_j(i(j))$ of \mathbf{x}_j for each $1 \leq j \leq q$ such that $\mathbf{x}_j(i(j))$ does not appear as a component of $\mathbf{x}_{j'}$ if $j' < j$.

The following claim is proved easily. (We omit the proof.)

Lemma 7.2. For an array $\mathfrak{X} \subset (\tau^{-n}(x))^p$ of p -tuples in $\tau^{-n}(x)$, we set

$$|\mathfrak{X}| := \{x' \in \tau^{-n}(x) \mid x' \text{ is a component of some } \mathbf{x} \in \mathfrak{X}\} \subset \tau^{-n}(x).$$

If $\#\mathfrak{X} > p(q-1)$, there is an independent array of q elements in \mathfrak{X} .

The next lemma explains the motivation for Definition 7.1.

Lemma 7.3. There exist $n_0 > 0$ (depending on q and hence on ε) such that, for any $\delta > 0$ and any $n \geq n_0$, we can find a family of smooth functions $g_k : S^1 \rightarrow \mathbb{R}$, $1 \leq k \leq K$, such that the following property holds for the family (84): For any

$x \in S^1 \setminus \text{Per}_\delta(\tau, n)$, any $m \geq 1$ and any array $\tilde{\mathfrak{X}} = (\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_1)$ of q elements in $(\tau^{-mn}(x))^p$ such that

$$\mathfrak{X} := \left(\mathbf{x}_j := (\tau^{(m-1)n} \tilde{\mathbf{x}}_j(i))_{i=1}^p \in (\tau^{-n}(x))^p \right)_{j=1, \dots, q}$$

is independent, we have

$$\det D\Phi_{x, \tilde{\mathfrak{X}}} |_Z \geq 1$$

for some q -dimensional subspace $Z \subset \mathbb{R}^K$.

Proof. Let us consider an arbitrary point $s \in S^1 \setminus \text{Per}_\delta(\tau, n)$. For $\rho > 0$, let $V_s(\rho)$ be the open ρ -neighborhood of s in S^1 . For $q \in \tau^{-n}(s)$, let $U_{s,q}(\rho)$ be the connected component of $\tau^{-n}(V_s(\rho))$ containing q , so that $\tau^{-n}(V_s(\rho))$ is the disjoint union of $U_{s,q}(\rho)$ for $q \in \tau^{-n}(s)$. Since $s \notin \text{Per}(\tau, n)$, we have $\tau^k(q_0) \neq q_1$ for any distinct $q_0, q_1 \in \tau^{-n}(s)$ and any $1 \leq k \leq n$. So we can choose $\rho(s) > 0$ so small that

$$(89) \quad \tau^k(U_{s,q_0}(\rho(s))) \cap U_{s,q_1}(\rho(s)) = \emptyset$$

for any $q_0, q_1 \in \tau^{-n}(s)$ and any $1 \leq k \leq n$. In particular, if $x \in \tau^{-n}(V_s(\rho(s)))$, we have $\tau^k(x) \notin \tau^{-n}(V_s(\rho(s)))$ for $1 \leq k \leq n$.

We take functions $g_{s,q} : S^1 \rightarrow \mathbb{R}$ for $q \in \tau^{-n}(s)$ so that $g_{s,q}$ is supported on $U_{s,q}(\rho(s))$ and satisfies

$$\frac{d}{dx} g_{s,q}(x) = 2\ell^n \quad \text{on } U_{s,q}(\rho(s)/3) \quad \text{and} \quad \left| \frac{d}{dx} g_{s,q}(x) \right| < 4\ell^n \quad \text{on } S^1.$$

By compactness, we can and do take a finite subset $H \subset S^1$ so that $V_s(\rho(s)/3)$ for $s \in H$ cover $S^1 \setminus \text{Per}_\delta(\tau, n)$. Finally we define g_k , $1 \leq k \leq K$, as a rearrangement of $g_{s,q}$ for $s \in H$ and $q \in \tau^{-n}(x)$.

We check that the conclusion of the lemma holds if we define the functions g_k , $1 \leq k \leq K$, as above and if n is sufficiently large. Suppose that $x \in S^1$ and arrays $\tilde{\mathfrak{X}}$ and \mathfrak{X} are given as in the statement of the lemma. Since \mathfrak{X} is independent, there is $1 \leq i(j) \leq p$ for $1 \leq j \leq q$ such that $\mathbf{x}_j(i(j))$ is not a component of $\mathbf{x}_{j'}$ if $j' < j$. We take $s \in S^1$ so that $x \in V_s(\rho(s)/3)$ and select $1 \leq k(j) \leq K$ for $1 \leq j \leq q$ so that $g_{k(j)}$ corresponds to $g_{s,q}$ for $q \in \tau^{-n}(s)$ such that $\mathbf{x}_j(i(j)) \in U_{s,q}(\rho(s)/3)$. Let Z be the q -dimensional subspace of \mathbb{R}^K that contains the $s_{k(j)}$ -axis for $1 \leq j \leq q$. Observe that $D\Phi_{x, \tilde{\mathfrak{X}}} |_Z$ is represented by the $q \times q$ matrix whose (j, j') -element is

$$M_{j,j'} = \sum_{i=1}^p \sum_{l=0}^{mn-1} \ell^{l-mn} \frac{d}{dx} g_{k(j')}(\tau^l(\tilde{\mathbf{x}}_j(i))).$$

We regard this matrix as the sum of $M^{(0)} = (M_{j,j'}^{(0)})_{j,j'}$ and $M^{(1)} = (M_{j,j'}^{(1)})_{j,j'}$ with

$$M_{j,j'}^{(0)} = \sum_{i=1}^p \sum_{l=(m-1)n}^{mn-1} \ell^{l-mn} \frac{d}{dx} g_{k(j')}(\tau^l(\tilde{\mathbf{x}}_j(i))) \quad \text{and} \quad M_{j,j'}^{(1)} = M_{j,j'} - M_{j,j'}^{(0)}.$$

From the disjoint property of the orbits of the supports of $g_{k(j)}$ that follows from (89) and from the assumption that \mathfrak{X} is independent, we observe that

- (a) $M^{(0)}$ is lower triangular in the sense that $M_{j,j'}^{(0)} = 0$ if $j' > j$,
- (b) the diagonal components of $M^{(0)}$ are $2k$ for some $1 \leq k \leq p$, while the other components are bounded by $2p$ in absolute value, and
- (c) $M^{(1)}$ is a $q \times q$ matrix whose elements are bounded by $4\ell^{-n}/(1 - \ell^{-n})$.

Hence if $n \geq n_0$ for some large n_0 depending on q (and p, ℓ), we always have

$$\det(D\Phi_{x, \tilde{\mathbf{x}}}|_Z) = \det(M^{(0)} + M^{(1)}) \geq 1.$$

This completes the proof. \square

In the following, we fix the family of functions g_i given in the lemma above.

7.2. The exceptional set. In this subsections, we investigate the situation where the roof function f does *not* belong to $\mathcal{G}(J, n, \varepsilon, \delta; p)$ and derive a few consequences. So let us suppose that there is an arbitrarily large $t > 0$ and a point $z_0 = (x_0, y_0) \in X_f$ with $x_0 \notin \text{Per}_\delta(\tau, n)$ and $\xi_0 \in [-\theta_0, \theta_0]$ such that, for any subset $\mathcal{E} \subset \tau^{-n}(x_0)$ with $\#\mathcal{E} \leq pq = p\lceil 10a/\varepsilon \rceil$, we have

$$(90) \quad \sum^* \frac{1}{W^r(\mathbf{w}, t; f)(\xi_0, 1)} \geq \exp((\max\{p \cdot h(f) - a, 0\} + p(b - a) + \varepsilon)t)$$

where the sum \sum^* is taken over $\mathbf{w} = (\mathbf{w}(1), \dots, \mathbf{w}(p)) \in B(z_0, t; J; f)^p$ such that $T_f^{s_n(z_0, \mathbf{w}(i); t)}(\mathbf{w}(i)) \notin \mathcal{E} \times \{0\}$ for $i = 1, 2, \dots, p$.

We begin with a few basic estimates (which hold in general). From the definition of $B(z_0, t; J; f)$, we have

$$e^{at} \leq E(w) = \ell^{k(z_0, w; t)} \leq e^{bt}, \quad \text{that is,} \quad \frac{at}{\log \ell} \leq k(z_0, w; t) \leq \frac{bt}{\log \ell}$$

for $w \in B(z_0, t; J; f)$, where $k(z_0, w; t)$ is that defined in (19). Hence, if we set

$$m := \left\lfloor \frac{at}{n \log \ell} \right\rfloor,$$

we have $mn \leq k(z_0, w; t)$ and

$$f^{(mn)}(T_f^{s_{mn}(z_0, w; t)}(w)) \leq t \leq \frac{(m+1)n \log \ell}{a} \quad \text{for } w \in B(z_0, t; J; f).$$

Note that, for each $x \in \tau^{-mn}(x_0)$, we have

$$(91) \quad \#\{w \in B(z_0, t; J; f) \mid T_f^{s_{mn}(z_0, w; t)}(w) = x\} \leq \ell^{\lfloor bt/\log \ell \rfloor - mn} \leq \ell^{n+1} e^{(b-a)t}.$$

For each $\mathbf{x} = (\mathbf{x}(i))_{i=1}^p \in (\tau^{-n}(x_0))^p$, let us set

$$\Delta^*(\mathbf{x}) = \sum_{\tilde{\mathbf{x}} \rightarrow \mathbf{x}} \left\langle \ell^{mn} |\xi_0 - \hat{S}(\tilde{\mathbf{x}}, mn; f)| \right\rangle^{-r}$$

where $\hat{S}(\tilde{\mathbf{x}}, mn; f)$ is that defined in (87) (with $\mathbf{s} = 0$) and the sum $\sum_{\tilde{\mathbf{x}} \rightarrow \mathbf{x}}$ is taken over those $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}(i))_{i=1}^p \in (\tau^{-mn}(x_0))^p$ satisfying

$$(92) \quad \tau^{(m-1)n}(\tilde{\mathbf{x}}(i)) = \mathbf{x}(i) \quad \text{and} \quad f^{(mn)}(\tilde{\mathbf{x}}(i)) \leq \frac{(m+1)n \log \ell}{a} \quad \text{for } 1 \leq i \leq p.$$

We claim that the assumption (90) implies

$$(93) \quad \sum_{\mathbf{x} \in (\tau^{-n}(x) \setminus \mathcal{E})^p} \Delta^*(\mathbf{x}) \geq \exp\left(\left(\max\{p \cdot h(f) - a, 0\} + \frac{\varepsilon}{2}\right)t\right)$$

for any subset $\mathcal{E} \subset \tau^{-n}(x_0)$ with $\#\mathcal{E} \leq pq$, provided that t is sufficiently large. To check this claim, let us consider the quantity

$$\Delta(\tilde{\mathbf{x}}) = \sum_{\tilde{\mathbf{w}} \rightarrow \tilde{\mathbf{x}}} \frac{1}{W^r(\tilde{\mathbf{w}}, t; f)(\xi_0, 2)} \quad \text{for } \tilde{\mathbf{x}} \in (\tau^{-mn}(x))^p$$

where the sum $\sum_{\tilde{\mathbf{w}} \rightarrow \tilde{\mathbf{x}}}$ is taken over $\tilde{\mathbf{w}} = (\tilde{\mathbf{w}}(i))_{i=1}^p \in B(z, t; J; f)^p$ such that

$$(94) \quad \tilde{\mathbf{x}}(i) = T^{s_{mn}(\tilde{\mathbf{w}}(i))}(\tilde{\mathbf{w}}(i)) \quad \text{for } 1 \leq i \leq p.$$

Then, from (91), we have

$$\Delta(\tilde{\mathbf{x}}) \leq C_0(\ell^{n+1} \exp((b-a)t))^p \left\langle \ell^{mn} |\xi_0 - \hat{S}(\tilde{\mathbf{x}}, mn; f)| \right\rangle^{-r}$$

with C_0 a constant depending only on r and θ_0 , because

$$\frac{1}{W^r(\tilde{\mathbf{w}}, t; f)(\xi_0, 2)} \leq C_0 \left\langle \ell^{mn} |\xi_0 - \hat{S}(\tilde{\mathbf{x}}, mn; f)| \right\rangle^{-r}$$

for $\tilde{\mathbf{w}} \in B(z, t; J; f)^p$ satisfying (94). Hence, for $\mathbf{x} \in (\tau^{-n}(x_0) \setminus \mathcal{E})^p$,

$$\sum_{\tilde{\mathbf{x}} \rightarrow \mathbf{x}} \Delta(\tilde{\mathbf{x}}) \leq C_0(\ell^{n+1} \exp((b-a)t))^p \Delta^*(\mathbf{x}).$$

If we take the sum of the left hand side above over $\mathbf{x} \in (\tau^{-n}(x_0) \setminus \mathcal{E})^p$, the total equals the left hand side of (90). Therefore we obtain the claim (93) provided that t is sufficiently large.

We next give a consequence of (93). Let us write \mathbf{y}_k , $1 \leq k \leq \ell^{pn}$, for the elements of $(\tau^{-n}(x_0))^p$ and suppose that they are sorted so that $\Delta^*(\mathbf{y}_k) \geq \Delta^*(\mathbf{y}_{k'})$ if $k \leq k'$. For $1 \leq k \leq \ell^{pn}$, let

$$Y_k = \{x \in \tau^{-n}(x_0) \mid x \text{ is a component of } \mathbf{y}_{k'} \text{ for some } k' \leq k\}.$$

Let k_* be the maximum of $1 \leq k \leq \ell^{pn}$ such that $\#Y_k \leq pq$. Set $\mathcal{E} = Y_{k_*}$ in (93). Then, since $\Delta^*(\mathbf{x}) \leq \Delta^*(\mathbf{y}_{k_*})$ for $\mathbf{x} \in (\tau^{-n}(x_0) \setminus \mathcal{E})^p$, we have that

$$\ell^{np} \cdot \Delta^*(\mathbf{y}_{k_*}) \geq \exp\left(\left(\max\{p \cdot h(f) - a, 0\} + \frac{\varepsilon}{2}\right)t\right).$$

This implies

$$(95) \quad \Delta^*(\mathbf{y}_k) \geq \frac{1}{\ell^{np}} \exp((\max\{p \cdot h(f) - a, 0\} + (\varepsilon/2))t) \quad \text{for } 1 \leq k \leq k_*.$$

Since $\#Y_{k_*} > p(q-1)$, we can choose an independent (ordered) array $(\mathbf{x}_k)_{k=1}^q$ from \mathbf{y}_k , $1 \leq k \leq k_*$, by using Lemma 7.2. In conclusion, we found an array $(\mathbf{x}_k)_{k=1}^q$ of q elements in $(\tau^{-n}(x_0))^p$ that is independent and that (95) holds with \mathbf{y}_k replaced by \mathbf{x}_j for $1 \leq j \leq q$.

Finally we reconsider about the choice of $x_0 \in S^1$ and $\xi_0 \in [-\theta_0, \theta_0]$. Recall that these are given from our assumption that the condition in the definition of $\mathcal{G}(J, n, \varepsilon, \delta; p)$ does not hold for f . But, by continuity, it is possible to shift these points a little so that they belong to some grids and that the conclusion of the argument above remains true for them (with slight difference in the constants). Precisely, for each $m > 0$, we choose a set $P(m)$ of points in $S^1 \times [-\theta_0, \theta_0]$ such that $\#P(m) \leq C_0 \ell^{2(1+\varepsilon)mn}$ and that the $\ell^{-(1+\varepsilon)mn}$ -neighborhood of those points cover $S^1 \times [-\theta_0, \theta_0]$. Then we can shift the point (x_0, ξ_0) to a nearby point in $P(m)$ so that the conclusion at the end of the last paragraph remains true.

Let us summarize the argument in this subsection as follows:

Lemma 7.4. *If $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$ does not belong to $\mathcal{G}(J, n, \varepsilon, \delta; p)$, we can find*

- (a) *an arbitrarily large integer $m \geq 1$,*
- (b) *a point $(x_0, \xi_0) \in P(m)$,*
- (c) *an independent array $(\mathbf{x}_k)_{k=1}^q$ of q elements in $(\tau^{-n}(x_0))^p$,*

such that

$$\sum_{\tilde{\mathbf{x}} \rightarrow \mathbf{x}_k} \left\langle \ell^{mn} |\xi_0 - \hat{S}(\tilde{\mathbf{x}}, mn; f)| \right\rangle^{-r} \geq \exp \left(\left(\max\{p \cdot h(f) - a, 0\} + \frac{\varepsilon}{2} \right) \frac{mn \log \ell}{a} \right)$$

where the sum $\sum_{\tilde{\mathbf{x}} \rightarrow \mathbf{x}_k}$ is taken over $\tilde{\mathbf{x}} \in (\tau^{-mn}(x_0))^p$ satisfying (92) with $\mathbf{x} = \mathbf{x}_k$.

7.3. The end of the proof. Let $\varepsilon, \delta > 0$ and $J = [a, b]$ are those given in the statement of Theorem 3.3. We take functions $g_k : S^1 \rightarrow \mathbb{R}$ for $1 \leq k \leq K$ as in Lemma 7.3 and consider the families (84) for all $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$. (The choice of $\sigma > 0$ will be given below.) For each of such families, we prove that $f_{\mathbf{s}}$ does *not* belong to $\mathcal{G}(J, n, \varepsilon, \delta; p)$ only when the parameter $\mathbf{s} \in R(\sigma)$ belongs to a subset with zero Lebesgue measure. This implies that the subset $\mathcal{G}(J, n, \varepsilon, \delta; p)$ is a prevalent subset.

Remark 7.5. For the last statement, recall Remark 3.4. The Lebesgue measure on the finite dimensional subspace of $C^\infty(S^1)$ spanned by g_k , $1 \leq k \leq K$, is the transverse measure to $\mathfrak{F}(y_{\min}, y_{\max}, \kappa_0) \setminus \mathcal{G}(J, n, \varepsilon, \delta; p)$.

Let $\eta > 0$ be a small real number that we will specify later. (At least, we suppose that η is much smaller than ε .) Then let $\sigma > 0$ be so small that

$$e^{-\eta} \cdot f(x) \leq f_{\mathbf{s}}(x) \leq e^{\eta} \cdot f(x) \quad \text{and} \quad |h(f_{\mathbf{s}}) - h(f)| < \eta$$

for $\mathbf{s} \in R(\sigma)$ and $f \in \mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$. For $x_0 \in S^1$ and $m \geq 1$, let $B(x_0, mn)$ be the set of points x in $\tau^{-mn}(x_0)$ satisfying

$$(96) \quad f^{(mn)}(x) \leq e^{\eta} \cdot \frac{mn \log \ell}{a}.$$

Note that, when m is sufficiently large, we have

$$\#B(x_0, mn) \leq \exp \left(h(f) \cdot e^{2\eta} \cdot \frac{mn \log \ell}{a} \right)$$

For an integer $m \geq 1$, a point $(x_0, \xi_0) \in P(m)$, an array $(\mathbf{x}_j)_{j=1}^q$ of q elements in $(\tau^{-n}(x_0))^p$ and an array $(\tilde{\mathbf{x}}_j)_{j=1}^q$ of q elements in $(\tau^{-mn}(x_0))^p$ such that

$$(97) \quad \tau^{(m-1)n}(\tilde{\mathbf{x}}_j(i)) = \mathbf{x}_j(i) \quad \text{for } 1 \leq i \leq p \text{ and } 1 \leq j \leq q,$$

we define the function

$$\Xi_m((x_0, \xi_0); (\mathbf{x}_j)_{j=1}^q; (\tilde{\mathbf{x}}_j)_{j=1}^q) : R(\sigma) \rightarrow \mathbb{R}$$

on the parameter space $R(\sigma)$ by

$$\Xi_m((x_0, \xi_0); (\mathbf{x}_j)_{j=1}^q; (\tilde{\mathbf{x}}_j)_{j=1}^q)(\mathbf{s}) = \prod_{j=1}^q \left\langle \ell^{mn} |\xi_0 - \hat{S}(\tilde{\mathbf{x}}_j, mn, f_{\mathbf{s}})| \right\rangle^{-r}.$$

If the array $(\mathbf{x}_j)_{j=1}^q$ is independent, we have from the choice of the functions g_i that

$$\int_{R(\sigma)} \Xi_m((x_0, \xi_0); (\mathbf{x}_j)_{j=1}^q; (\tilde{\mathbf{x}}_j)_{j=1}^q)(\mathbf{s}) d\mathbf{s} \leq C_0 \ell^{-mnq}$$

for a constant C_0 depending only on r . Therefore we have

$$(98) \quad \sum^{**} \int_{R(\sigma)} \Xi_m((x_0, \xi_0); (\mathbf{x}_j)_{j=1}^q; (\tilde{\mathbf{x}}_j)_{j=1}^q)(\mathbf{s}) d\mathbf{s} \\ \leq C_0 \ell^{-mnq} \cdot \ell^{2(1+\varepsilon)mn} \cdot \exp \left(pq \cdot h(f) \cdot e^{2\eta} \cdot \frac{mn \log \ell}{a} \right)$$

for sufficiently large m , where the sum \sum^{**} is taken over combinations of

- a point $(x_0, \xi_0) \in P(m)$,
- an independent array $(\mathbf{x}_j)_{j=1}^q$ of q elements in $\tau^{-n}(x_0)^p$ and
- an array $(\tilde{\mathbf{x}}_j)_{j=1}^q$ in $(B(x_0, mn))^p \subset \tau^{-mn}(x_0)^p$ satisfying (97).

Let $\mathcal{X} \subset R(\sigma)$ be the set of parameters $\mathbf{s} \in R(\sigma)$ such that $f_{\mathbf{s}}$ belongs to $\mathfrak{F}(y_{\min}, y_{\max}, \kappa_0)$ and does not satisfy the condition in the definition of $\mathcal{G}(J, n, \varepsilon, \delta; p)$. From the conclusion in the last subsection given in Lemma 7.4, we see that

$$\mathcal{X} \subset \limsup_{m \rightarrow \infty} \mathcal{X}_m$$

where \mathcal{X}_m is the set of parameters $\mathbf{s} \in R(\sigma)$ such that

$$\sum^{**} \Xi_m((x_0, \xi_0); (\mathbf{x}_j)_{j=1}^q; (\tilde{\mathbf{x}}_j)_{j=1}^q)(\mathbf{s}) \geq \exp \left(q \left(p \cdot e^{-\eta} \cdot h(f) - a + \frac{\varepsilon}{2} \right) \frac{mn \log \ell}{a} \right)$$

Comparing this with (98), we see that the Lebesgue measure of \mathcal{X}_m is bounded by

$$C_0 \exp \left(\left(\frac{2a(1+\varepsilon)}{q} + (e^{2\eta} - e^{-\eta})p \cdot h(f) - \frac{\varepsilon}{2} \right) \cdot \frac{qmn \log \ell}{a} \right).$$

From the choice of q in (86), we can take small $\eta > 0$ (and also $\sigma > 0$ accordingly) so that this bound decreases exponentially with respect to m . Hence Lebesgue measure of \mathcal{X} is zero by Borel-Cantelli lemma.

REFERENCES

- [1] V. Baladi and M. Tsujii. Dynamical determinants and spectrum for hyperbolic diffeomorphisms. In *Geometric and probabilistic structures in dynamics*, volume 469 of *Contemp. Math.*, pages 29–68. Amer. Math. Soc., Providence, RI, 2008.
- [2] P. Buser. *Geometry and spectra of compact Riemann surfaces*, volume 106 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1992.
- [3] P. Giulietti, C. Liverani, and M. Pollicott. Anosov flows and dynamical zeta functions. *Ann. of Math. (2)*, 178(2):687–773, 2013.
- [4] I. Gohberg, S. Goldberg, and N. Krupnik. *Traces and determinants of linear operators*, volume 116 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2000.
- [5] B. R. Hunt, T. Sauer, and J. A. Yorke. Prevalence: a translation-invariant “almost every” on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)*, 27(2):217–238, 1992.
- [6] B. R. Hunt, T. Sauer, and J. A. Yorke. Prevalence. An addendum to: “Prevalence: a translation-invariant ‘almost every’ on infinite-dimensional spaces” [Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 2, 217–238; MR1161274 (93k:28018)]. *Bull. Amer. Math. Soc. (N.S.)*, 28(2):306–307, 1993.
- [7] Z.-h. Lu. Topological pressure of continuous flows without fixed points. *J. Math. Anal. Appl.*, 311(2):703–714, 2005.
- [8] R. Mañé. *Ergodic theory and differentiable dynamics*, volume 8 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1987. Translated from the Portuguese by Silvio Levy.
- [9] W. Ott and J. A. Yorke. Prevalence. *Bull. Amer. Math. Soc. (N.S.)*, 42(3):263–290 (electronic), 2005.
- [10] W. Parry and M. Pollicott. An analogue of the prime number theorem for closed orbits of Axiom A flows. *Ann. of Math. (2)*, 118(3):573–591, 1983.
- [11] M. Pollicott. On the mixing of Axiom A attracting flows and a conjecture of Ruelle. *Ergodic Theory Dynam. Systems*, 19(2):535–548, 1999.
- [12] M. Pollicott and R. Sharp. Exponential error terms for growth functions on negatively curved surfaces. *Amer. J. Math.*, 120(5):1019–1042, 1998.
- [13] D. Ruelle. Locating resonances for Axiom A dynamical systems. *J. Statist. Phys.*, 44(3-4):281–292, 1986.

- [14] L. Stoyanov. Ruelle transfer operators for contact anosov flows and decay of correlations. 2013.
- [15] H. Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.
- [16] M. Tsujii. A measure on the space of smooth mappings and dynamical system theory. *J. Math. Soc. Japan*, 44(3):415–425, 1992.
- [17] M. Tsujii. Decay of correlations in suspension semi-flows of angle-multiplying maps. *Ergodic Theory Dynam. Systems*, 28(1):291–317, 2008.
- [18] M. Tsujii. Quasi-compactness of transfer operators for contact Anosov flows. *Nonlinearity*, 23(7):1495–1545, 2010.
- [19] M. Tsujii. Contact Anosov flows and the Fourier-Bros-Iagolnitzer transform. *Ergodic Theory Dynam. Systems*, 32(6):2083–2118, 2012.

DEPARTMENT OF MATHEMATICS, KYUSHU UNIVERSITY, MOTOOKA 744, NISHI-KU, FUKUOKA,
819-0395, JAPAN

E-mail address: tsujii@math.kyushu-u.ac.jp